

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international(43) Date de la publication internationale
23 septembre 2004 (23.09.2004)

PCT

(10) Numéro de publication internationale
WO 2004/080378 A2(51) Classification internationale des brevets⁷ : A61K(21) Numéro de la demande internationale :
PCT/FR2004/000499

(22) Date de dépôt international : 3 mars 2004 (03.03.2004)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
03/02633 4 mars 2003 (04.03.2003) FR(71) Déposant (pour tous les États désignés sauf US) : UNI-
VERSITE DE NICE - SOPHIA ANTIPOLIS [FR/FR];
Grand Château, 28 Avenue Valrose, BP 2135, F-06103
Nice Cedex 2 (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : BAR-
CELLINI-COUGET, Sylvie [FR/FR]; 1853 avenue de
Pibonson, Le Mas des Roses, F-06250 Mougins (FR).
DOGLIO, Alain [FR/FR]; 3341 route de l'Abadie,
F-06730 Saint Andre (FR).(74) Mandataires : BECKER, Philippe etc.; Becker et Asso-
ciés, 35, rue des Mathurins, F-75008 Paris (FR).(81) États désignés (sauf indication contraire, pour tout titre de
protection nationale disponible) : AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,
CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,
KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,
MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,
PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.(84) États désignés (sauf indication contraire, pour tout titre de
protection régionale disponible) : ARIPO (BW, GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasi-
en (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT,
BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR,
HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR),
OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,
MR, NE, SN, TD, TG).

Déclaration en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv) pour US
seulement

Publiée :

— sans rapport de recherche internationale, sera republiée
dès réception de ce rapportEn ce qui concerne les codes à deux lettres et autres abrégia-
tions, se référer aux "Notes explicatives relatives aux codes et
abréviations" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.

(54) Title: METHODS AND TOOLS FOR SCREENING ACTIVE RNA IN CELLULO

(54) Titre : METHODES ET OUTILS POUR LE CRIBLAGE D'ARN ACTIFS IN CELLULO

(57) Abstract: The invention relates to methods and compositions for screening active RNA in cellulose. The invention more partic-
ularly relates to methods for the preparation and selection of random RNA providing a cell with a desired phenotype, random RNA
banks and the production thereof, in addition to cellular populations which are transformed by said banks. The invention further
relates to active RNA identified for the purpose of research or for therapeutical reasons in order to induce a desired phenotype in a
cell in vitro, ex vivo or in vivo, more particularly in human cells.(57) Abrégé : La présente invention concerne des méthodes et compositions pour le criblage d'ARN actifs in cellulose. Elle concerne
notamment des procédés de préparation et de sélection d'ARN aléatoires conférant à une cellule un phénotype désiré, des banques
d'ARN aléatoires et leur production, ainsi que des populations cellulaires transformées par ces banques. Elle concerne également
l'utilisation des ARN actifs identifiés à des fins de recherche ou thérapeutiques, pour induire un phénotype désiré dans une cellule in
vitro, ex vivo ou in vivo, notamment dans des cellules humaines.

WO 2004/080378 A2

BEST AVAILABLE COPY

Méthodes et Outils pour le criblage d'ARN actifs in cellulo

Introduction

5 La présente invention concerne des méthodes et compositions pour le criblage et la sélection d'ARN actifs in cellulo. Elle concerne notamment des procédés de préparation et de production de banques de cassettes d'expression d'ARN ou de banques de cellules codant des ARN aléatoires, ainsi que leurs utilisations pour la sélection d'ARN actifs capables de produire ou altérer un phénotype cellulaire d'intérêt. L'invention concerne également
10 l'utilisation des ARN actifs identifiés à des fins de découverte de nouveaux gènes (« drug discovery »), de validation de la fonction de gènes, de développement d'outils pharmacologiques, de diagnostic ou à des fins thérapeutiques avec notamment la possibilité de corriger, grâce aux ARN, l'expression de phénotypes pathogènes dans une cellule in vitro, ex vivo ou in vivo, particulièrement dans des cellules humaines et animales.

15

Domaine de l'invention

La possibilité d'agir ou d'interférer de manière spécifique avec des cibles biologiques peut avoir de multiples applications, notamment dans les domaines thérapeutique, diagnostique,
20 vaccinal ou expérimental. Ainsi, la capacité de réguler l'expression d'un gène peut permettre de bloquer ou de restaurer une activité dans une cellule et/ou de corriger une pathologie. La faculté de bloquer l'expression de gènes de pathogènes peut permettre de stopper leur développement, etc.

L'expression d'un gène résulte de la superposition de plusieurs étapes comprenant la
25 synthèse de l'ARN messenger, le métabolisme intracellulaire de cet ARN, la production de la protéine et enfin la stabilité et l'activité de cette protéine. Inhiber l'expression d'un gène consiste donc à agir sur une de ces étapes. Par ailleurs, d'autres métabolites, voies de signalisation ou composants cellulaires peuvent être ciblés comme des sucres, lipides, nucléosides, etc. Dans cette perspective, les ARN apparaissent comme de puissants outils
30 capables d'agir efficacement, spécifiquement et de façon discriminante sur la plupart des étapes de l'expression génique ou sur des cibles biologiques. En effet, la plasticité

structurale de l'ARN génère une diversité de motifs structuraux capables de se lier avec la plupart des cibles biologiques, et notamment de molécules organiques présentes dans les cellules (ARN, ADN, protéines et divers métabolites tels que lipides, sucres, etc). Les interactions mises en jeu sont généralement très spécifiques, ce qui confère à l'ARN une forte sélectivité vis-à-vis de sa cible, et l'accumulation intracellulaire d'ARN n'est ni
5 toxique, ni immunogène.

La possibilité de développer, sélectionner et exploiter le potentiel des ARN permettrait donc d'envisager de nombreuses approches thérapeutiques avantageuses, car sélectives et non
10 toxiques pour l'organisme (Famulok and Verma, 2002). Différentes stratégies ont été envisagées dans l'art antérieur pour identifier ou fabriquer de tels ARN. On peut citer notamment la technologie SELEX, destinée à fabriquer et sélectionner in vitro des ARN structuraux aléatoires (Gold, 1995). On peut également citer des approches visant à exprimer des banques d'ARN antisens ou de ribozymes dans des cellules, afin de tester leur
15 activité biologique (WO99/41371, WO99/32618, WO98/32880). Néanmoins, à ce jour, aucune approche de l'art antérieur ne permet de produire ou de sélectionner des ARN actifs in cellulo, dans des conditions optimales et dans une configuration active. Ainsi, les approches in vitro ne sont pas prédictives des propriétés des molécules in vivo, car elles ne permettent pas de présager de leur pénétration cellulaire, de leur localisation cellulaire, de
20 leur stabilité, de leur résistance aux nucléases, et de leur activité biologique. De même, les approches décrites dans les demandes citées précédemment ne sont pas directement exploitables pour l'expression d'ARN structuraux, qui sont actifs dans des configurations particulières, ou pour l'expression efficace d'ARN actifs dans les cellules. De même, la demande n° WO00/58455 propose des approches pour le criblage d'ARN, mais dont les
25 conditions de mise en œuvre ne permettent pas une exploitation efficace du potentiel de ces agents. Ainsi, les outils, vecteurs et constructions mentionnés dans cette demande impliquent des étapes répétées de sélection in vitro, imposent une sélection individuelle in cellulo des constructions, et ne permettent donc pas une sélection rapide et simple de banques d'ARN dans des configurations actives dans les cellules.

Résumé de l'Invention

La présente demande permet de remédier aux inconvénients de l'art antérieur. La présente demande fournit maintenant de nouvelles méthodes et de nouveaux outils pour la
5 production, l'expression et la sélection dans les cellules, notamment de mammifères, de séquences d'ARN dites « actives ».

La présente invention réside notamment dans la mise au point de conditions particulières de préparation et de criblage des ARN actifs, permettant de sélectionner des composés
10 candidats particulièrement avantageux. L'invention repose notamment sur l'utilisation de constructions particulières, permettant aux séquences ARN d'être exprimées et localisées à l'intérieur de la cellule, dans une conformation active et stable, et dans des compartiments choisis. L'invention repose également sur la conception et l'utilisation de cassettes d'expression particulières assurant une grande efficacité d'expression dans les cellules de
15 mammifères et ayant un caractère inductible. L'invention décrit en outre la construction de vecteurs viraux améliorés, permettant une sélection efficace et à grande échelle des ARN actifs, et la production de produits directement utilisables pour des applications thérapeutiques. Les motifs actifs ARN sont ainsi sélectionnés directement dans la cellule, à l'aide d'approches innovantes, et notamment selon l'approche désignée par l'appellation
20 «SECAR» («*Selective Enrichment of Cellular Aptamer RNA*»).

Un premier aspect de l'invention concerne donc des procédés de sélection, identification ou optimisation in cellulo d'ARN actifs. Les procédés de l'invention peuvent être adaptés pour la sélection in cellulo d'ARN actifs capables de conférer à une cellule un phénotype désiré,
25 y compris d'ARN actifs capables d'altérer l'activité d'une cible biologique déterminée. Les procédés de l'invention comprennent, plus particulièrement :

- 1) la mise en contact d'une banque d'acides nucléiques comprenant une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression dérivée d'un gène VA d'un adénovirus, éventuellement rendue inductible,
30 exprimant un ARN structural aléatoire distinct, ou d'une partie de celle-ci, avec une

population de cellules dans des conditions permettant l'infection desdites cellules par lesdits rétrovirus recombinants,

- 2) la sélection des cellules possédant le phénotype désiré, et
- 3) l'identification de la ou des cassette(s) contenue(s) dans la ou lesdites cellules, ou de l'ARN qu'elles expriment, ledit ARN étant capable de conférer le phénotype désiré à ladite population de cellules.

L'insertion d'un motif actif au sein d'une structure ARN stable (ARN VA) permet de concevoir une entité moléculaire globale, désignée par les inventeurs par le terme "aptARN", composée d'une partie commune à tous les aptARN (la navette VA) qui sert à stabiliser, présenter et transporter le motif actif dans la cellule. Ce type de véhicule intracellulaire pour aptamère est innovant et permet d'améliorer considérablement la production intracellulaire d'aptamères ARN par rapport aux techniques de l'art antérieur.

- 15 La possibilité d'induire l'expression des cassettes au cours des étapes de sélection permet de valider directement l'effet de l'ARN actif dans les cellules sélectionnées grâce à la simple comparaison des phénotypes cellulaires observés en comparant les états réprimés et induits (figure 15, panneau A). Ce procédé simplifie donc considérablement les étapes de criblage phénotypique sur cellules de mammifères pour l'identification in cellulo d'ARN actifs, il constitue de ce fait une amélioration notable de ce type d'approche par rapport à l'art antérieur.

La production et la sélection directes, dans les cellules, des ARN actifs stabilisés selon l'invention est avantageuse car elle permet de sélectionner des molécules se trouvant déjà dans une conformation active intracellulaire.

Comme il sera décrit en détails dans la suite du texte, les étapes 1) à 3) peuvent être répétées une ou plusieurs fois à partir des cassettes ou ARN actifs identifiés à l'étape 3), afin d'améliorer, au fur et à mesure des cycles, la sélection et/ou la qualité des ARN actifs, et/ou d'adapter ou de moduler leurs propriétés.

- Par ailleurs, selon l'application envisagée, la banque d'acides nucléiques de départ peut être plus ou moins complexe et plus ou moins contrainte. Ainsi, il peut s'agir d'une banque aléatoire non contrainte, ou d'une banque produite à partir de la séquence d'un gène cible donné ou comprenant des motifs ou des résidus imposés, selon le profil des ARN actifs
- 5 recherchés. Il peut également s'agir d'une banque pré-sélectionnée in vitro, notamment d'une banque codant des ARN pré-sélectionnés in vitro, par exemple d'une banque produite à partir de la séquence d'ARN sélectionnés in vitro pour une propriété particulière (par exemple leur capacité de liaison à une cible d'intérêt).
- 10 Par rapport aux approches existantes du type SELEX, l'avantage d'un procédé est de sélectionner, non pas un motif structural ARN considéré isolément en dehors du contexte de son expression intracellulaire, mais une entité moléculaire globale, l'aptARN, présentant une affinité pour une cible pré-identifiée et déjà dans une conformation conforme à celle de son expression intracellulaire.
- 15 D'autre part, à l'issue de l'étape 3) du dernier cycle réalisé (ou éventuellement d'un ou plusieurs ou de tout cycle intermédiaire), une étape supplémentaire 4) peut être mise en œuvre, afin de confirmer l'activité biologique du ou des ARN actifs sélectionnés.
- 20 De manière avantageuse, l'invention propose donc une combinaison innovante d'éléments génétiques particuliers, selon une méthodologie de sélection particulière, formant ainsi un concept global, intégré et simplifié de criblage phénotypique. Par rapport aux techniques proposées dans l'art antérieur, les outils et méthodes décrites dans la présente demande présentent de nombreux avantages. Ainsi, l'invention montre que des cassettes d'expression
- 25 dérivées d'un gène VA d'adénovirus peuvent être intégrées et exprimées dans un contexte de vecteur rétroviral. Le transfert rétroviral est très efficace et permet de maîtriser le nombre de copies intégrées par cellules, et les cassettes construites, dérivées de gènes VA forment des cassette intégrées, c'est-à-dire contenant toutes les informations nécessaires à une transcription efficace (promoteur, signal Stop, structure responsable de la stabilité de l'ARN
- 30 dans la cellule, structure responsable de l'export des ARN vers le cytoplasme, possibilités diverses d'insertion de séquences exogènes et tolérance assez large par rapport au promoteur). L'invention montre en outre que la structure génétique des constructions peut

être modifiée pour diriger et déterminer la localisation intracellulaire de l'ARN produit (noyau, cytoplasme) et pour rendre l'expression inductible de manière simple. La combinaison du transfert rétroviral et d'une cassette dérivée d'un gène VA permet donc d'obtenir des niveaux d'expression élevés et contrôlés dans la plupart des lignées cellulaires et d'utiliser des constructions d'ARN identiques in vitro et in cellulo, assurant ainsi une sélection plus efficace de molécules actives et une validation plus rapide des cassettes actives identifiées.

L'invention est applicable dans de nombreux domaines, et notamment pour valider la fonction d'un gène, rechercher de nouvelles cibles impliquées dans une fonction cellulaire, trouver et produire de nouvelles molécules pour le diagnostic, la pharmacologie ou la thérapeutique, etc.

Un autre aspect de l'invention réside dans des banques d'acides nucléiques aléatoires et/ou actifs, éventuellement contenus dans des cassettes et/ou vecteurs. Ainsi, un objet particulier de l'invention concerne une banque d'acides nucléiques caractérisée en ce qu'elle comprend une pluralité d'espèces de virus recombinants, chaque espèce de virus comprenant une cassette d'expression exprimant un ARN aléatoire (et/ou actif) distinct sous contrôle d'un promoteur transcrit par l'ARN polymérase III, en particulier dérivé de la séquence d'un gène VA d'adénovirus. L'ARN peut être un ARN structural aléatoire ou de séquence définie. De préférence, les virus sont des rétrovirus recombinants.

Un autre objet de l'invention réside dans des cassettes d'expression d'ARN actifs, comprenant une séquence codant ledit ARN insérée dans un promoteur dérivé d'un gène VA d'un adénovirus, ledit promoteur pouvant comprendre en outre une séquence conférant un caractère inductible et/ou une modification qui permet de retenir l'ARN dans le noyau.

L'invention concerne également tout vecteur comprenant une telle cassette d'expression, ainsi que des cellules recombinantes les contenant.

L'invention a encore pour objet toute composition comprenant un ARN actif, une cassette, un vecteur ou une cellule tels que définis ci-avant et/ou identifiés ou produits par le procédé de sélection décrit dans l'invention.

- 5 L'invention a encore pour objet des compositions pharmaceutiques comprenant un ARN actif, une cassette, un vecteur ou une cellule tels que définis ci-avant, et un véhicule ou excipient acceptable sur le plan pharmaceutique.

L'invention concerne également une composition pharmaceutique caractérisée en ce qu'elle
10 comprend un ARN actif, ledit ARN actif comprenant une séquence active insérée dans un ARN VA modifié, ledit ARN VA modifié comprenant un motif ARN assurant sa localisation dans le noyau de la cellule et/ou une séquence conférant un caractère inductible.

L'invention concerne également une composition pharmaceutique caractérisée en ce qu'elle
15 comprend la séquence ARN active ou le motif actif identifié au sein de l'ARN actif. Cette séquence ARN active ou ce motif actif ARN isolé peut en outre être ultérieurement modifié chimiquement par exemple de façon à améliorer sa stabilité en solution.

L'invention concerne également des outils, constructions, lignées, etc., utiles pour la
20 production des compositions définies ci-dessus, notamment des gènes VA modifiés, des gènes tRNA modifiés, des cassettes U6 modifiées, des vecteurs ou cellules les comprenant, et leurs utilisations.

L'invention concerne encore des procédés de production de compositions pharmaceutiques,
25 comprenant (i) le criblage d'une banque d'ARN aléatoires comme décrit ci-avant, permettant d'obtenir une cassette d'expression d'un ARN actif et (ii) le conditionnement de la cassette ou de la séquence ARN active dans tout excipient ou véhicule acceptable sur le plan pharmaceutique.

30 L'invention est utile pour l'identification, la production, l'expression et/ou la sélection de tout ARN actif sur des cellules, notamment de mammifères. Elle permet la préparation de

composés actifs dans différentes situations, notamment pour la production d'agents thérapeutiques, en particulier anti-infectieux, anti-cancéreux, agissant sur le métabolisme de la cellule, agissant sur les processus de différenciation de la cellule, sur la capacité de croissance de la cellule, etc. Comme indiqué précédemment, l'invention offre de nombreux avantages par rapports à l'art antérieur, et permet notamment la sélection directe, rapide et simple de molécule d'ARN actives dans les cellules.

Légende des Figures

10 La figure 1 décrit le gène VA1 de l'adénovirus de type 2 ainsi que son produit de transcription.

La figure 1A représente les séquences nucléotidiques du gène VA1-RNA (région transcrite) (SEQ ID NO : 1). Les nucléotides notés en gras, de 13 à 24 et de 59 à 68, représentant respectivement la boîte A et la boîte B, sont les éléments nécessaires à la transcription du gène VA1 par l'ARN polymérase de type III. Les nucléotides 157 à 160 (TTTT), notés en gras, représentent le signal d'arrêt de transcription de la polymérase III. Les nucléotides 93 à 118 soulignés, correspondant à la séquence de la boucle IV, sont délétés dans la construction VAAIV (SEQ ID NO : 2).

20 La figure 1B représente la structure secondaire de l'ARN VA1 de l'Adénovirus de type 2, obtenue d'après le logiciel Mac DNASIS Pro V3.6. La partie délétée dans l'ARN VAAIV est indiquée par deux flèches (nucléotides 93 à 118). Le nucléotide 120 est muté dans la construction VAAIV (C est muté en T).

25 La figure 2 représente l'étude comparative de la production et de la localisation cellulaire des ARN VA1, VAAIVSrf et VATAR*.

Figure 2A : Séquence TAR* (SEQ ID NO : 23) insérée dans le site de restriction SrfI du gène VAAIVSrf pour générer le gène VATAR*. L'oligonucléotide TAR* est dérivé de la séquence TAR du Virus de l'Immunodéficience Humaine (HIV-1) (Yamamoto et al., 2000).

Figure 2B : Analyse de l'expression des ARN VA1, VAAIVSrf et VATAR* par Northern Blot. Les différents ARN sont extraits des cellules 293, 48h après transfection de ces cellules par des plasmides contenant les gènes VA1, VAAIVSrf et VATAR*.

Figure 2C : Etude de la localisation cellulaire des ARN VA1, VAAIVSrf et
5 VATAR* dans les cellules humaines 293. Les ARN sont visualisés par hybridation in situ, 48h après transfection de ces cellules par des plasmides contenant les gènes VA1, VAAIVSrf et VATAR*.

La figure 3 représente le promoteur du gène U6 snRNA humain (Genebank access :
10 X07425) . Les éléments requis pour le recrutement et l'activité de l'ARN polymérase de type III sont le DSE (Distal Sequence Element : nucléotides -221 à -216), le PSE (Proximal Sequence Element : nucléotides -64 à -46) et la boîte TATA (TATA Box : nucléotides -31 à -24). Au niveau du site +1 de transcription se trouve le site de restriction Sal I (GTCGAC), le premier nucléotide transcrit étant le premier nucléotide de ce site (G).

15

La figure 4 représente la structure secondaire de l'ARN VAAIV obtenue avec le logiciel Mac DNASIS Pro V3.6.

La figure 5 représente la structure secondaire de l'ARN VAAIVSrf obtenue avec le logiciel
20 Mac DNASIS Pro V3.6. Le site de clonage SrfI, qui permet l'insertion des séquences exogènes dans l'ARN VAAIV Srf, est indiqué par une flèche.

La figure 6 décrit les caractéristiques de l'ARN nVAAIVSrf (localisation nucléaire).

Figure 6A : Séquence du gène codant l'ARN nVAAIVSrf (SEQ ID NO : 4). Les
25 nucléotides notés en gras, de 13 à 24 et de 59 à 68, représentant respectivement la boîte A et la boîte B, sont les éléments nécessaires à la transcription du gène VA1 par l'ARN polymérase de type III. Les nucléotides 138 à 141 (TTTT), notés en gras, représentent le signal d'arrêt de transcription de la polymérase III. Le site de restriction Srf I (93 à 100) apparaît souligné. Les nucléotides 120, 121 et 122 écrits en minuscule représentent les
30 mutations introduites dans l'ARN VAAIVSrf pour altérer l'hélice terminale et conférer à cet ARN une localisation nucléaire.

Figure 6B : Schéma de la structure secondaire de l'ARN nVA Δ IVSrf obtenue avec le logiciel DNASIS Pro V3.6. Le site de clonage SrfI, qui permet l'insertion des séquences exogènes dans l'ARN nVA Δ IV Srf, est indiqué par une flèche.

Figure 6C : Etude de la localisation cellulaire des ARN nVA Δ IVSrf. Les ARN sont visualisés par hybridation in situ, 48h après transfection des cellules 293 par un plasmide contenant le gène VA Δ IVSrf.

La figure 7 présente les différents gènes VA inductibles (VAi) : leurs séquences, les séquences des différents oligonucléotides ADN qui ont été utilisés pour leur construction ainsi que l'étude de leur niveau d'expression cellulaire.

Figure 7A : Séquences des gènes VAI Δ O (SEQ ID NO : 15), OVAi (SEQ ID NO : 16) et OVAi Δ O (SEQ ID NO : 17). La numérotation des séquences s'effectue à partir du point +1 de transcription ; les séquences situées en amont sont numérotées négativement. Les séquences des Boîtes A et B, ainsi que le signal d'arrêt de transcription, sont notées en gras. La séquence opératrice TetO1 est soulignée ; le site de clonage SrfI est noté en italique.

Figure 7B : Séquences des oligonucléotides ayant servi à la construction des gènes VAI par réaction d'élongation d'amorce ou par réaction de polymérisation en chaîne. Les séquences des Boîtes A et B ainsi que le signal d'arrêt de transcription sont notées en gras. Les séquences opératrices TetO1 sont soulignées ; les sites de clonage SrfI et PvuII sont notés en italique. Les régions complémentaires entre les oligonucléotides VAI up (SEQ ID NO : 18) et VAI down (SEQ ID NO : 19) apparaissent en lettres minuscules. Les zones d'hybridation des oligonucléotides VAI PvuII (VAi PvuII 5' (SEQ ID NO : 20), VAI PvuII 3' (SEQ ID NO : 22) ou OVAi PvuII 5' (SEQ ID NO : 21)) avec les régions codantes des gènes VAI sont également représentées en minuscule.

Figure 7C : Analyse de l'expression des ARN VAI dans les cellules Hela T-Rex après induction par la doxycycline (Dox). Les différents gènes VAI (OVAi, VAI Δ O et OVAi Δ O) ont été clonés dans le site NheI du plasmide pBabe (voir Figure 10). Les cellules de la lignée d'encapsulation 293GP ont été utilisées pour la production des rétrovirus recombinants BabeVAi. Puis les cellules Hela T-Rex (Invitrogen ref: R714-07) qui expriment le répresseur TetR ont été transduites par les différents rétrovirus recombinants

babeVAi. Après sélection des cellules transduites par la puromycine (gène de sélection de pBabe), les différents gènes VAi ont été activés grâce à l'ajout de Doxycycline (Dox : 1,5µg/ml). Les ARN ont été extraits aux temps indiqués et l'expression des ARN VAi a été analysée par Northern blot.

5

La figure 3 représente la structure secondaire de l'ARN VAiO obtenue avec le logiciel Mac DNASIS Pro V3.6. La région correspondant à la séquence tetO1 est limitée par des flèches.

La figure 9 représente la structure secondaire des ARN h9U6, h20U6 et nh20U6 obtenue avec le programme Mac DNASIS Pro V3.6. Les ARN schématisés sur cette figure représentent les structures de base invariantes de ces ARN (partie navette). Le motif actif de l'ARN est cloné dans le site Sfr I indiqué par une flèche.

La figure 10 représente le vecteur rétroviral pBabe (Morgenstern et Land, 1990). Le site Nhe I correspond au site de clonage des différentes cassettes d'expression des ARN actifs.

La figure 11 montre la relation qui peut exister entre le niveau d'expression d'un ARN actif et son activité biologique. A titre d'exemple, l'ARN actif utilisé ici est l'aptamère TAR* décrit dans la figure 2B. Ce motif TAR* correspond à un motif ARN capable de réprimer la réplication du virus de l'immunodéficience humaine (HIV). Ce motif TAR* a été cloné dans le site Sfr I de la cassette VAAIVSfr. La construction VATAR* a été insérée dans le site de clonage NheI du vecteur rétroviral pBabe (voir figure 10). Les cellules de la lignée d'encapsidation 293GP ont été utilisées pour la production des retrovirus recombinants Babe et Babe/VATAR*.

L'efficacité anti-HIV de VATAR* a été mesurée en utilisant les cellules indicatrices P4 qui constituent un système couramment utilisé pour étudier la multiplication du HIV (Charneau et al., 1994). Les cellules P4 ont été transduites par les rétrovirus Babe et Babe/VATAR* et sélectionnées par l'ajout de puromycine. Le niveau d'expression de l'ARN VATAR* a été analysé par Northern-blot au sein de la population de cellules P4 sélectionnée (population parentale), ainsi que dans 2 clones cellulaires indépendants (clones 1 et 4) isolés à partir de cette population. L'infection de ces différents systèmes cellulaires par le HIV-1 démontre

30

que le clone 1 exprimant un important niveau d'ARN VATAR* résiste plus efficacement à la multiplication du HIV. Le taux d'infection par le HIV-1 est estimé grâce à la mesure de l'activité du gène rapporteur β -galactosidase contenue dans les cellules P4 et les résultats sont présentés en pourcentage du maximum d'infection.

5

La Figure 12 décrit une application du procédé de sélection d'ARN actifs à partir d'une librairie de cassettes d'expression d'ARN aléatoires clonée dans le vecteur rétroviral pBabe. Dans cette application, les ARN actifs ont été sélectionnés pour leur capacité à rendre des cellules Hela résistantes à l'apoptose induite par la Staurosporine.

10

Figure 12A : Les différentes étapes de synthèse de la librairie de séquences aléatoires. La librairie d'oligonucléotides ADN simple brin comprend de 5' vers 3' : 8 bases fixes surlignées (run A), 26 bases aléatoires puis 8 bases fixes surlignées (run B). L'oligonucléotide run A, complémentaire à Run B est hybridé avec la librairie d'oligonucléotides ADN simple brin et allongé grâce à l'ADN polymérase fragment de

15 Kleenow. Cet ADN double brin néosynthétisé est appelé librairie d'ADN double brin aléatoire.

20

Figure 12 B : Librairie de vecteurs Babe VA aléatoire. La librairie de vecteurs rétroviraux recombinants Babe VA aléatoire est générée par clonage de la librairie d'ADN double brin aléatoire dans le site Srf I du gène VAAIV Srf. Cette librairie est donc constituée d'une collection de cassettes d'expression d'ARN VAAIV Srf contenant chacun un motif aléatoire et cloné dans le vecteur rétroviral pBabe.

25

Figure 12 C : Procédé de sélection des clones de cellules Hela résistantes à la staurosporine. La librairie de vecteurs Babe VA aléatoire est transfectée dans les cellules d'encapsulation de la lignée 293GP pour produire la librairie de rétrovirus VA aléatoires. Cette librairie rétrovirale est utilisée pour transduire la librairie de cassettes aléatoires dans les cellules de la lignée Hela de façon à ce que chaque cellule transduite contienne en moyenne une seule cassette VA aléatoire. Après sélection des cellules transduites, l'apoptose cellulaire est induite par ajout de staurosporine (0,8 μ M – 6h). Les cellules résistantes à la staurosporine sont amplifiées sous forme de clone cellulaire et les cassettes

d'expression des ARN actifs présentes dans chacun des clones sont analysées. A ce stade, 12 cassettes ont été sélectionnées et leur activité anti-apoptotique a été étudiée (Figure 12D).

Figure 12 D : Validation de l'activité anti-apoptotique des ARN identifiés grâce au procédé de sélection SECAR. Les cassettes d'expression d'ARN actifs sélectionnés (en Figure 12C) ont été clonées dans le plasmide pBabe. Ces plasmides recombinants ont été utilisés pour produire des rétrovirus recombinants grâce à la transfection des cellules de la lignée d'encapsulation 293GP. Deux lignées cellulaires différentes (Hela et Jurkat) ont été transduites indépendamment avec les différents surnageants rétroviraux (clone 2, 5, 9, 11, 13, 14, 15 16, C, J, L, et N) et sélectionnées par la puromycine. Les ARN cellulaires totaux
10 issus de ces différentes populations de cellules transduites ont été extraits, puis analysés par Northern blot.

Figure 12 E : Validation individuelle de l'activité anti-apoptotique des ARN identifiés grâce au procédé SECAR. Les populations de cellules Hela et Jurkat exprimant les ARN actifs sélectionnés en Figure 12C ont été soumises à des tests de résistance à la staurosporine (0,8µM pendant 6heures).
15

Cellules Hela : Les résultats présentés dans l'histogramme reflètent le nombre de cellules résistantes à la staurosporine (Unité arbitraire).

Cellules Jurkat : Les résultats présentés dans l'histogramme correspondent au pourcentage de cellules résistantes à la staurosporine (Mesure de la perméabilité membranaire mitochondriale).
20

La Figure 13 illustre le procédé de l'invention avec une librairie contrainte pré-sélectionnée in vitro (iv SECAR). Utilisation des cassettes d'expression inductibles.

Figure 13 A : Construction d'une librairie de cassettes de transcription VAiO aléatoire.
25

Etape 1 : les oligonucléotides banque Sens et banque iv Antisens sont hybridés puis allongés par l'ADN polymérase fragment de Kleenow pour générer la librairie de fragment VAiO aléatoire.

Etape 2 : la librairie de cassettes d'expression VAiO aléatoire est obtenue par réaction de polymérisation en chaîne grâce à l'utilisation d'une Taq DNA polymérase en utilisant pour
30

matrice la librairie de fragment VAiO aléatoire et les oligonucléotides sens VAI5' et antisens VAI end Banque.

Etape 3 : la librairie de cassettes de transcription T7 VAiO aléatoire est obtenue par réaction de polymérisation en chaîne grâce à l'utilisation d'une Taq DNA polymérase en utilisant
5 pour matrice la librairie de cassettes de transcription VAiO aléatoire et les oligonucléotides VApT7 en 5' et VA end en 3'.

Figure 13 B : Production d'ARN dérivés du gène VA : ARN VAAIV Srf, ARN VATAR*, librairie d'ARN VAiO aléatoire. Les ARN sont synthétisés in vitro grâce à l'utilisation de l'ADN polymérase du bactériophage T7, puis sont visualisés sur gel
10 d'agarose après coloration au Bromure d'Éthidium (200 ng par piste).

Figure 13 C : Obtention des cassettes de transcription dérivées du gène VA à partir des ARN synthétisés in vitro. Une réaction de RT-PCR est réalisée à partir de différents substrats ARN synthétisés in vitro en utilisant les oligonucléotides VapT7 et VA end. Les cassettes de transcription correspondant aux différents ARN VA ont été obtenues :

- 15 cassette d'expression VAAIV Srf obtenue à partir des ARN VAAIV Srf;
- cassette d'expression VATAR* obtenue à partir des ARN VATAR* ;
- cassette d'expression VAiO aléatoire obtenue à partir des ARN VAiO aléatoire. De même un mélange équimoléculaire des trois types d'ARN (VAAIV Srf, VATAR* et VAiO aléatoire) permet l'obtention d'un mélange équimoléculaire des cassettes correspondantes.
- 20 Les produits des réactions de RT-PCR sont visualisés sur gel d'agarose 2% après coloration au Bromure d'Éthidium.

La figure 14 fournit un exemple de construction de gène hybride U6/VA.

Il s'agit ici de mettre le gène VAiO sous le contrôle du promoteur du gène murin
25 U6 : gène mU6/VAiO.

La première étape a consisté à introduire le promoteur du gène U6 murin dans le site NheI du plasmide pBabe. Le promoteur du gène U6 murin (mU6) a été obtenu par réaction de polymérisation en chaîne en utilisant les amorces mU6 amont et mU6 aval. Dans l'amorce mU6 aval, quatre bases ont été ajoutées (en italique, gras et souligné). L'ajout de ces bases a
30 permis de générer le site PmeI indiqué dans la construction pBabe mU6/VAiO. Dans

l'amorce mU6 aval, les quatre bases AAAC soulignées correspondent aux quatre dernières bases du promoteur U6 et aux quatre premières bases du site de restriction Pme I (GTTTAAAC). Dans la réaction de polymérisation en chaîne, la matrice utilisée correspond à l'ADN génomique de cellules murines. Le produit de cette réaction de polymérisation en chaîne a été purifié, puis ligué dans le plasmide pBabe digéré par Nhe I. Le produit de ligation est appelé pBabe mU6.

La deuxième étape a consisté à insérer le gène VAIo dans le site PmeI du plasmide pBabe mU6. Le gène VAIo a été produit par une réaction de polymérisation en chaîne utilisant les amorces VAIo Start et VAIo End NheI, la matrice étant le plasmide pBabe VAIo.

10 L'amorce VAIo End NheI contient le site de restriction NheI noté en gras et souligné. Le produit de cette réaction a été purifié puis ligué dans le plasmide pBabe mU6 digéré par PmeI. Le produit de ligation est appelé pBabe mU6/VAIo.

La figure 15 fournit une description générale des stratégies de criblage d'ARN actifs selon l'invention.

Panneau A : Procédé de sélection des cassettes d'expression d'ARN actif inductibles (SECAR) : Ce procédé permet la sélection in cellulo d'ARN capables de conférer à une cellule un phénotype donné. Le criblage phénotypique présenté dans ce panneau est effectué grâce à l'utilisation de cassettes d'expression inductibles ; il s'effectue en six étapes.

20 Étape 1 : La librairie de cassettes d'expression d'ARN aléatoires est transférée sur les cellules d'intérêt, via un vecteur rétroviral. Le transfert est ainsi optimisé : il est adapté à tout type cellulaire et permet de contrôler le nombre de cassettes insérées par cellule.

Étape 2 : Les cellules ayant intégré le nouveau matériel génétique (cassette d'expression d'ARN aléatoire) sont sélectionnées grâce à un gène de sélection présent dans le vecteur rétroviral.

25 Étape 3 : Grâce à l'utilisation de cassettes d'expression inductibles, l'expression des cassettes est induite afin d'activer la transcription des ARN aléatoires.

Étape 4 : La sélection des cellules qui expriment les ARN actifs et qui présentent le phénotype d'intérêt est effectuée. Ces cellules sont amplifiées.

Étape 5 (facultative) : Les cassettes d'expression des ARN actifs présentes dans les cellules sélectionnées à cette étape peuvent être copiées et à nouveau transférées dans les cellules d'intérêt. Ainsi les étapes 1 à 4 peuvent être répétées autant que nécessaire.

Étape 6 : Dans les clones cellulaires obtenus en fin d'étape 4 ou 5, l'activité des cassettes
5 est validée par comparaison des états induits ou réprimés. L'ARN est validé en tant qu'ARN actif quand le phénotype cellulaire recherché est observé dans la seule condition pour laquelle la cassette d'expression est induite. Au contraire, l'ARN n'est pas validé en tant qu'ARN actif quand le phénotype cellulaire recherché est observé quelles que soient les conditions : cassette d'expression induite ou cassette d'expression réprimée.

10

Panneau B : Procédé de sélection des cassettes d'expression d'ARN actif après enrichissement préalable en séquences actives (ivSECAR pour «*in vitro Selective Enrichment of Cellular Aptamer RNA*»)

Ce procédé, comme le procédé SECAR, permet la sélection in cellulo d'ARN capables de
15 conférer à une cellule un phénotype donné. Dans ce cas, le criblage est effectué en utilisant une librairie d'ARN aléatoires, enrichie en ARN capables de lier une cible donnée. Les premières étapes du criblage sont effectuées in vitro, à partir d'une banque d'ARN aléatoires ; les ARN utilisés pour le criblage in vitro ont une structure identique à ceux qui seront exprimés in cellulo dans les étapes ultérieures.

20 Étape 1 : A partir d'une librairie de cassettes de transcription d'ARN aléatoires, la librairie d'ARN aléatoire est produite in vitro. Les ARN ainsi synthétisés in vitro ont les mêmes propriétés structurales que les ARN qui sont synthétisés in cellulo, à partir d'une banque de cassettes d'expression d'ARN aléatoires.

Étape 2 : La banque d'ARN aléatoire est mise en contact avec la cible d'intérêt.

25 Étape 3 : Les ARN capables de lier la cible sont retenus grâce à un procédé approprié.

Étape 4 : Les ARN retenus en fin d'étape 3 sont utilisés pour produire une nouvelle librairie de cassette de transcription d'ARN aléatoires, enrichie en séquences actives. Les étapes 1 à 4 sont ainsi renouvelées à volonté. En fin d'étape 4, une librairie « restreinte » de cassettes d'expression d'ARN actifs est obtenue.

30 Étape 5 : Cette librairie restreinte est par la suite utilisée dans des essais cellulaires selon le procédé SECAR décrit dans le panneau A.

Description Détaillée de l'Invention

5 L'invention concerne, de manière générale, des outils et procédés de sélection in cellulo d'ARN actifs capables de conférer à une cellule un phénotype désiré, ainsi que l'utilisation de ces ARN actifs ou de tout ADN codant dans le domaine expérimental ou pharmaceutique, par exemple. Comme indiqué précédemment, l'invention utilise, de manière particulièrement avantageuse, la combinaison de vecteurs rétroviraux et de
10 cassettes dérivées de gènes VA, permettant un criblage simple, efficace et prédictif, in vitro comme in vivo.

ARN actifs

15 La présente invention est adaptée à la production, l'expression et/ou la sélection de toute molécule d'ARN active, c'est-à-dire de molécules d'ARN capables d'interagir avec et/ou d'altérer l'activité de composants biologiques, et/ou de conférer à une cellule un phénotype donné.

Le terme ARN actif inclut notamment des ARN structuraux (tels que les aptamères) et des
20 ARN non-structuraux, tels que des antisens, des ribozymes ou des ARN interférants (siARN, miARN ou leurs précurseurs). Les ARN actifs peuvent être de longueur variable, comprise typiquement entre 8 et 500 bases, plus préférentiellement entre 8 et 200, notamment entre 8 et 150. Les ARN actifs sont généralement synthétisés sous forme de molécules simple-brin, même s'ils peuvent ultérieurement adopter des structures tri-
25 dimensionnelles de type boucle, régions double-brins, hélices, etc.

Dans un mode spécifique et préféré de l'invention, l'ARN actif est un ARN structural, notamment un aptamère (Famulok and Verma, 2002) (Hermann and Patel, 2000). Les aptamères sont des oligonucléotides capables de lier spécifiquement une molécule cible.

Dans un autre mode spécifique de l'invention, l'ARN actif est un ARN interférant (siARN, mi ARN ou leurs précurseurs) (McManus and Sharp, 2002) (Scherr et al., 2003) (Famulok and Verma, 2002) ou un ARN antisens.

La présente demande est en effet particulièrement adaptée à la conception et au criblage
5 d'ARN dont l'activité requiert une configuration spatiale particulière.

Les procédés décrits dans la présente demande permettent de sélectionner des ARN actifs à partir de collections ou banques de séquences aléatoires, générales ou restreintes, pouvant comprendre un très grand nombre de séquences aléatoires distinctes. Comme il sera décrit
10 plus en détails dans la suite du texte, les séquences aléatoires peuvent être toute molécule d'ADN ou d'ARN comprenant au moins un élément de séquence non connu, plus précisément dont une partie au moins de la séquence est aléatoire. La sélection des ARN actifs selon l'invention est réalisée in cellulo, grâce à l'insertion des séquences aléatoires dans des cassettes d'expression et dans des conditions particulières. Un des avantages
15 principaux de la sélection in cellulo est que la cible de l'ARN actif n'est pas forcément choisie à priori ; de plus l'ARN actif sélectionné est réellement efficace dans les conditions d'utilisation cellulaire.

Promoteur dérivé du gène VA d'un adénovirus

20

Comme indiqué ci-avant, une caractéristique de l'invention réside dans l'utilisation de cassettes d'expression particulières, permettant la production in cellulo d'ARN actifs, dans des configurations optimales.

25 En particulier, d'une manière particulièrement avantageuse, les constructions, compositions et méthodes de l'invention utilisent un promoteur dérivé de la séquence d'un gène VA d'adénovirus.

Plusieurs travaux, dont ceux des inventeurs, ont démontré l'utilité du gène VA1 pour
30 exprimer dans la cellule des ribozymes, des aptamères ou des antisens, dont la séquence

était parfaitement définie (Medina and Joshi, 1999) (Barcellini et al., 1998) (Bertrand et al., 1999) (Cagnon et al., 1995) (Gwizdek et al., 2001).

La présente demande démontre pour la première fois la possibilité et l'efficacité d'utiliser ce
 5 type de construction pour la production et le criblage de banques d'ARN aléatoires (ou de
 banques contraintes) in cellulo, notamment dans un contexte rétroviral. Elle décrit
 également différentes modifications dans le gène VA afin de générer des cassettes
 améliorées conférant un meilleur contrôle des paramètres d'expression. Les séquences
 codant pour les ARN actifs peuvent ainsi être clonées dans la cassette, principalement dans
 10 le domaine central du gène. Le positionnement de la séquence active au sein du domaine
 central ne modifie pas le niveau de production, et la localisation de l'ARN chimérique dans
 la cellule (figure 2). L'utilisation d'un tel promoteur est particulièrement avantageuse pour
 l'expression d'ARN actifs structuraux, notamment d'aptamères.

15 Le génome des adénovirus contient deux petits gènes qui sont transcrits par l'ARN
 polymérase III, les gènes VA1 et VA2 (Mathews and Shenk, 1991). L'ARN VA1, et à un
 moindre degré l'ARN VA2, sont abondamment produits pendant le cycle de réplication de
 l'adénovirus. L'interaction entre l'ARN VA1 et une kinase cellulaire (PKR ou p68 kinase)
 bloque l'effet antiviral induit par les interférons et facilite la production de nouveaux virus
 20 (Mathews and Shenk, 1991).

Le gène VA1 des adénovirus code pour un ARN court (ARN VA1 160nt) caractérisé par
 une structure secondaire riche en région double brin. L'organisation génétique de ce gène
 transcrit par l'ARN polymérase III est simple et comprend une courte région promoteur en
 25 position intragénique (boîte A et boîte B) et un signal d'arrêt de transcription (voir figure
 1A). La structure secondaire de l'ARN VA1 est bien connue (figure 1B). La séquence du
 gène VA1 de l'adénovirus de type 2 est représentée sur la figure 1A (SEQ ID NO :1). La
 séquence de l'ARN codé correspondant est donnée ci-dessous.

30 1- GGGCACUCUU CCGUGGUCUG GUGGAUAAU UCGCAAGGGU
 AUCAUGGCGG ACGACCGGGG UUCGAACCCC GGAUCCGGCC
 GUCCGCCGUG AUCCAUGCGG UUACCGCCCG CGUGUCGAAC

CCAGGUGUGC GACGUCAGAC AACGGGGGAG CGCUCCUUUU -160

Les boîtes A et B correspondent aux nucléotides 11-24 et 59-68 en gras, respectivement, et le domaine central contenant la boucle IV correspond aux nucléotides 93 à 118.

5

Dans un premier mode de réalisation particulier de l'invention, la cassette comprend la séquence d'un gène VA1 d'adénovirus délété de tout ou une partie fonctionnelle de la boucle IV, dans laquelle la séquence codant l'ARN actif est insérée. Comme indiqué ci-dessus, le domaine central contenant la boucle IV correspond aux nucléotides 93 à 118 du gène VA. La séquence d'ARN transcrits par des cassettes d'expression selon l'invention comprenant la séquence d'un gène VA délété de la boucle IV est fournie ci-après :

SEQ ID NO : 2 (VAΔIV):

15 1- GGGCACUCUU **CCGUGGUCUG** **GUGGAUAAAU** UCGCAAGGGU
 AUCAUGGCGG ACGACCGGGG **UUCGAACCCC** GGAUCCGGCC
 GUCCGCCGUG AUAUCCAGGU GUGCGACGUC AGACAACGGG
 GGAGCGCUCC UUUU -134

20 La séquence SEQ ID NO : 2 (VAΔIV) est dérivée de VA1 Ad2 par délétion du domaine central de 93 à 118. Le nucléotide 94 (120 dans VA1) a été muté de façon à créer un site de clonage EcoRV (en italique souligné). Le site de coupure EcoRV se situe entre 92 et 93.

SEQ ID NO : 3 (VAΔIVSrf)

25 1- GGGCACUCUU **CCGUGGUCUG** **GUGGAUAAAU** UCGCAAGGGU
 AUCAUGGCGG ACGACCGGGG **UUCGAACCCC** GGAUCCGGCC
 GUCCGCCGUG AUGCCCGGGC AUCCAGGUGU GCGACGUCAG
 ACAACGGGGG AGCGCUCCUU UU -142 (SEQ ID NO : 3).

30 La séquence SEQ ID NO : 3 (VAΔIVSrf), ci-dessus, est dérivée de VAΔIV après insertion de la séquence du site SrfI dans le site EcoRV (en italique souligné). Le site de la coupure SrfI se situe entre 96 et 97.

- Dans un autre mode de réalisation particulier, qui peut être combiné avec le précédent, la double-hélice terminale du gène VA1 est altérée. Cette altération permet de contrôler la localisation cellulaire de l'ARN synthétisé (Gwizdek et al., 2001). Ainsi, lorsque l'hélice terminale du gène VA1 est intacte, l'ARN synthétisé se trouve de façon naturelle essentiellement dans le cytoplasme. Cette localisation est adaptée pour cribler des ARN actifs sur des cibles biologiques présentes essentiellement dans ce compartiment cellulaire. De manière alternative, lorsque l'hélice terminale du gène VA1 est altérée, l'ARN synthétisé se trouve essentiellement localisé dans le noyau. Cette localisation est adaptée pour cribler des ARN actifs sur des cibles biologiques essentiellement nucléaires. L'hélice terminale est formée essentiellement par l'appariement entre les 20 premiers et les 20 derniers nucléotides de la séquence de l'ARN codé par le gène VA1. L'altération de l'hélice peut être obtenue par différentes modifications introduites au sein de ces éléments de séquence, notamment par mutation, délétion et/ou insertion, de préférence par mutation. Lorsque la séquence de l'hélice est inférieure à environ 15 bases, l'hélice est altérée fonctionnellement et l'ARN VA1 est retenu dans le noyau. Egalement, lorsque la double hélice est mutée de manière à introduire une succession de nucléotides non appariés qui forment une ouverture au sein de l'hélice (bulle), cette dernière est altérée fonctionnellement et l'ARN est retenu dans le noyau.
- 20 La séquence d'un ARN transcrit par une cassette d'expression selon l'invention comprenant la séquence d'un gène VA délété de la boucle IV et une double-hélice terminale altérée est fournie ci-après (SEQ ID NO : 4) :

25 1- GGGCACUCUU **CCGUGGUCUG** GUGGAUAAAU UCGCAAGGGU
 AUCAUGGCGG ACGACCGGGG **UUCGAACCCC** GGAUCCGGCC
 GUCCGCCGUG AUGCCCGGGC AUCCAGGUGU GCGACGUCAa
 uaAACGGGGG AGCGCCCUUU U -141

- La séquence SEQ ID NO : 4 (nVA Δ IVSrf) est dérivée de VA Δ IVSrf par mutation des nucléotides 120 à 122 (minuscules) et délétion du nucléotide 136.
- 30

Dans un mode de réalisation particulièrement préféré, qui peut être combiné avec l'un ou les précédents, le gène VA1 comprend une séquence conférant un caractère inductible. Quelle que soit l'utilisation faite de l'ARN actif (antisens, ribozyme, aptamère, siARN, miARN et leurs précurseurs) exprimé au sein d'une cellule, le contrôle de son expression est un
5 élément déterminant. Dans cet objectif, la présente invention montre, pour la première fois, que la structure des gènes VA1 peut être modifiée de manière à rendre leur expression inductible, tout en conservant par ailleurs les caractéristiques décrites précédemment (niveau d'expression élevé, contrôle de localisation, possibilité d'insertion de séquences actives dans le domaine central). Ces résultats sont particulièrement inattendus compte tenu
10 de la structure intragénique du promoteur VA, et permettent la conception de cassettes d'expression ayant des propriétés remarquables pour le criblage d'ARN actifs in cellulo (voir Figure 15A). Ainsi, l'invention propose et permet, pour la première fois, la validation in cellulo de l'activité des cassettes sélectionnées par comparaison des états induits et réprimés. Cette approche permet de réduire le nombre de faux positifs et de simplifier et d'accélérer
15 l'identification des ARN d'intérêt.

Plus particulièrement, la présente invention montre qu'un gène VA inductible peut être construit par insertion d'une ou plusieurs séquences conférant un caractère inductible, typiquement entre les boîtes A et B du gène et/ou en amont du gène, de préférence en
20 remplacement de tout ou partie des séquences normalement présentes. Dans le cas où la séquence est insérée entre les boîtes A et B, cette nouvelle séquence se retrouve dans une région de l'ARN VA1 qui est transcrite, et provoque donc des altérations de la structure secondaire native de l'ARN VA1 susceptibles de le rendre moins stable. Afin de corriger ces altérations, des mutations compensatrices peuvent être introduites au sein de la séquence
25 VA1, de manière à corriger ces défauts de structure et à recréer ainsi les appariements naturels de l'ARN VA1. Ces cassettes optimisées permettent de contrôler l'expression des ARN actifs dans les cellules, et d'améliorer les conditions d'un criblage, ou de l'utilisation en recherche ou en thérapeutique.

30 Dans un mode particulier, l'invention concerne donc un acide nucléique comprenant la séquence d'un gène VA d'un adénovirus, modifié par insertion d'une ou plusieurs

séquences conférant un caractère inductible, de préférence entre les boîtes A et B et/ou en amont du gène, plus préférentiellement en remplacement de tout ou partie des séquences natives.

- 5 La séquence conférant un caractère inductible peut être toute séquence conférant une sensibilité à un facteur agissant en trans. Il peut s'agir de préférence d'un site de liaison d'un facteur transcriptionnel ou d'un répresseur, ou de tout autre agent ou molécule. Dans un mode de réalisation préféré, il s'agit d'une ou plusieurs séquences opératrices d'un promoteur bactérien régulé, par exemple du promoteur de la tétracycline.

10

- Le promoteur des gènes bactériens tet contient deux types de séquences opératrices O1 et O2 qui servent de site de fixation pour le répresseur TetR (Hillen et Berens, 1994). Chacun des sites tetO1 et tetO2 lie un homodimère de TetR. Des études ont montré que le site de fixation tetO2 avait une affinité pour l'homodimère TetR trois à cinq fois plus importante
15 que tetO1 (Hillen et Berens, 1994). Les séquences opératrices tetO1 ont ainsi été insérées à l'intérieur de séquences promotrices hétérologues eucaryotes afin d'obtenir un système d'expression eucaryote regulable par la tétracycline (Gossen and Bujard, 1992). Cependant, une telle modification ou application n'a jamais été envisagée à partir du gène VA1 des adénovirus, dont la manipulation est particulièrement délicate compte tenu de la présence du
20 promoteur dans les séquences transcrites.

- Nous avons inséré une ou deux séquences tetO1 en amont du gène VA1 afin de rendre son expression inductible par ajout de tétracycline. Ainsi, les séquences situées entre les Boîte A et B ou placées en amont du premier nucléotide transcrit du gène VA1 ont été remplacées
25 par des séquences tetO1. En absence de tétracycline, l'expression du gène VA1 modifié peut ainsi être réprimée par la liaison des homodimères de TetR sur les séquences cis-régulatrices. L'ajout de tétracycline permet de libérer le répresseur de sa fixation sur les séquences tetO, et au gène VA1 d'être normalement exprimé.

- 30 Dans un mode particulier, l'invention concerne donc un acide nucléique comprenant la séquence du gène VA1 d'un adénovirus, modifié par insertion d'une ou plusieurs séquences

opératrices servant de site de fixation pour le répresseur TetR. De préférence, la ou les séquences opératrices sont insérées entre les boîtes A et B de la séquence du gène VA et/ou placées en amont du gène, plus préférentiellement en remplacement de tout ou partie des séquences natives. L'invention concerne également toute cassette d'expression comprenant
5 une séquence codant un ARN actif insérée sous contrôle d'un promoteur VA inductible, (VAi) notamment par la tétracycline. La séquence d'exemples spécifiques de gènes VA inductibles construits selon l'invention est représentée dans les séquences SEQ ID NO: 15 (VaiO), SEQ ID NO: 16 (OVAi) et SEQ ID NO: 17 (OVAiO).

10 Des cassettes spécifiques dérivées de la séquence de gènes VA et permettant l'expression d'ARN aléatoires ou actifs sont décrites dans les exemples, qui représentent des objets particuliers de l'invention. On peut ainsi citer les cassettes VAAIVSrf, nVAAIVSrf, VAi et nVAi.

15 D'autre part, dans un mode de mise en œuvre particulier, les cassettes d'expression de l'invention comprennent, en plus du promoteur dérivé du gène VA d'un adénovirus, un second promoteur transcriptionnel, différent du promoteur VA. L'utilisation de telles constructions hybrides est notamment intéressante lorsque le promoteur dérivé du gène VA n'est pas fonctionnel dans un type cellulaire particulier. Dans ce cas, l'ARN actif et
20 configuré dans la séquence VA peut être exprimé sous contrôle d'un second promoteur, qui peut être de nature et d'origine variées. Avantagusement, le second promoteur est un promoteur transcrit par l'ARN polymérase de type III, et qui est fonctionnel dans le type cellulaire concerné, notamment un promoteur extra-génique. En aval de ce second promoteur, au niveau du site +1 de transcription, est clonée la cassette d'expression dérivée
25 du gène VA (VAAIV, VAAIV Srf, nVAAIV, VAi ou toute autre cassette dérivée du gène VA). Selon un mode préféré, le promoteur du gène U6 est choisi en tant que second promoteur permettant une transcription active du gène VA ainsi cloné.

Un objet particulier de l'invention réside ainsi dans une cassette d'expression hybride
30 comprenant une cassette d'expression dérivée d'un gène VA d'un adénovirus, éventuellement rendue inductible, exprimant un ARN actif (aptamère, antisens, ribozyme,

ARN interférent : Si ARN, miARN ou leurs précurseurs), placée en aval d'un second promoteur transcriptionnel. La construction de ce type de cassettes hybrides est détaillée dans les exemples (voir exemple A-3) et dans les figures.

5 Promoteurs dépendants de l'ARN Polymérase III

Dans certains modes de réalisation particuliers de l'invention, par exemple pour les cassettes hybrides, il peut être possible d'utiliser d'autres promoteurs dépendants de l'ARN Polymérase III pour l'expression des ARN actifs. Il peut s'agir d'un promoteur localisé à
10 l'intérieur des séquences transcrites (intragénique), comme le promoteur des gènes tRNA. Il peut également s'agir d'un promoteur localisé en amont des séquences transcrites (extragénique), comme les promoteurs du type U6 codant le petit ARN nucléaire U6 (snRNAU6). En outre, le promoteur utilisé peut être rendu inductible.

15 La transcription du gène U6 humain (hU6) génère un petit ARN nucléaire (ARN snU6) intégré dans un complexe ribo-nucléoprotéique responsable de l'épissage des ARN (Gmeiner, 2002). Le promoteur hU6 est placé en amont de la séquence transcrite, et les séquences correspondant à ce promoteur ne sont donc pas produites dans la cellule. Cela a pour avantage principal de n'imposer que peu de contraintes de séquence dans la région
20 transcrite de la cassette (Bertrand et al., 1997). En effet, le deuxième nucléotide transcrit correspond au premier nucléotide apporté par la séquence exogène et les derniers nucléotides correspondent au signal stop (figure 3). Il est donc possible, grâce à ce type de cassette, de produire dans la cellule des ARN artificiels (antisens, ribozymes, aptamères, ARN interférents : siARN, miARN et leurs précurseurs) dont on contrôle la quasi-totalité de
25 la séquence transcrite.

Afin de contrôler la stabilité des ARN générés ainsi que leur localisation cellulaire, différentes cassettes d'expression utilisant un promoteur U6 ont été construites par les inventeurs. Ces cassettes permettent l'expression d'ARN cytoplasmiques ou nucléaires,
30 structuraux ou non structuraux, contraints ou non-contraints. Des modifications ont ainsi été apportées au promoteur U6, dans le but de faciliter le clonage des séquences actives, de

maîtriser les séquences transcrites dès le point +1 de transcription, de stabiliser les séquences actives et de maîtriser la localisation intracellulaire de ces séquences.

Dans un premier mode de réalisation, une séquence servant de base structurale pour l'ARN
5 actif est placée en aval du promoteur U6. Cette séquence de structuration contient plus
préférentiellement une séquence capable de générer une (courte) hélice ARN (plus ou moins
stable et plus ou moins longue), et/ou un site de clonage bord franc et/ou un signal d'arrêt de
transcription (par exemple TTTTT). La séquence permettant la formation d'une hélice ARN
comprend typiquement deux régions complémentaire espacées par une région charnière. Ce
10 type de cassette est surtout utile pour l'expression intracellulaire de motifs d'ARN
structuraux. Lorsque le motif actif est lui-même structuré sous la forme d'une hélice, la
séquence de structuration comprend de préférence un motif navette court (typiquement de 5
à 12 bases). Lorsque le motif actif n'est pas structuré en hélice ou très court, la séquence de
structuration comprend de préférence un motif navette long, comprenant typiquement de 12
15 à 20 bases. Ce motif peut être structuré pour former une hélice parfaite assurant à l'ARN
produit une localisation cytoplasmique ou bien être structuré pour former une hélice
perturbée afin d'obtenir une localisation nucléaire de l'ARN généré.

Dans un autre mode de réalisation, qui peut être combiné au précédent, une séquence
20 formant une courte structure en épingle à cheveux est placée en aval du promoteur U6 (et,
le cas échéant, de la séquence de structuration). Cette séquence forme une structure ARN
stable placée à l'extrémité 3' de l'ARN actif ou aléatoire synthétisé, protégeant ainsi l'ARN
produit de la dégradation par des RNases. En aval de cette structure de stabilisation se
trouve le signal d'arrêt de transcription de la polymérase III : TTTTT.

25

Dans un mode de réalisation alternatif, la cassette comprend le promoteur U6 auquel est
directement liée la séquence codant l'ARN actif. Une telle cassette permet d'exprimer des
ARN dont la séquence est choisie dès le point +1 de transcription.

30 Selon un mode particulier, la cassette comprend un promoteur U6 tel que défini ci-dessus et
une ou plusieurs séquences conférant un caractère inductible, comme décrites

précédemment. La ou les séquences conférant le caractère inductible peuvent être placées par exemple en amont de la séquence du promoteur.

Des cassette spécifiques dérivées du promoteur U6 et permettant l'expression d'ARN
5 aléatoires ou actifs sont décrites dans les exemples, qui représentent des objets particuliers de l'invention. On peut ainsi citer les cassettes U6hélices, U6Srf, U6Tt, etc.

Dans d'autres modes de mise en œuvre particuliers, le promoteur transcrit par l'ARN
polymérase III est issu d'un promoteur des gènes ARNt. En outre, ces promoteurs peuvent
10 être modifiés pour leur conférer un caractère inductible.

Vecteurs

Les cassettes d'expression selon l'invention sont typiquement clonées ou comprises dans
15 des vecteurs, de clonage et/ou d'expression. De tels vecteurs peuvent être de nature et/ou d'origine variée, comme des plasmides, virus recombinants, vecteurs viraux, épisomes, chromosomes artificiels, phages, etc. De préférence, les vecteurs sont de type plasmide, viral ou épisomiques.

20 Ainsi, un objet particulier de l'invention réside dans un vecteur comprenant au moins une cassette d'expression telle que définie précédemment.

Lorsque le vecteur est de type plasmide, il peut être issu de différents plasmides connus ou commerciaux. Il comporte typiquement une origine de répllication compatible avec
25 l'utilisation désirée. Il peut comprendre en outre un gène de sélection et, éventuellement, une séquence d'intégration dans le chromosome. Des exemples de plasmides utilisables pour la production de vecteurs de l'invention sont notamment les plasmides pUC, pcDNA, pVV2, plasmides à répllication épisomique dérivés du virus Epstein-Bar du type OriP.

30 Dans un mode de mise en œuvre préféré, le vecteur est de type viral. Il peut s'agir d'un virus recombinant, c'est-à-dire d'une particule virale recombinante comprenant un génome

viral recombinant dans lequel est insérée au moins une cassette telle que définie précédemment. Il peut également s'agir d'un vecteur viral, c'est-à-dire d'une construction génétique comprenant un génome viral recombinant dans lequel est insérée au moins une cassette telle que définie précédemment.

5

Dans un mode de réalisation particulièrement préféré, le vecteur est un vecteur rétroviral ou un rétrovirus recombinant. La présente demande montre en effet que le transfert d'une cassette de l'invention dans des cellules cibles est très efficace après clonage de cette cassette dans un vecteur rétroviral. La présente demande montre notamment de manière
10 inattendue que les cassettes dérivées des gènes VAI1 peuvent être transduites efficacement dans les cellules cibles, dans un contexte rétroviral. En particulier, après intégration du vecteur rétroviral contenant la cassette dans le génome de ces cellules, le niveau d'expression de l'ARN chimérique est important et les propriétés de la cassette sont conservées.

15

Les virus ont été utilisés pour vectoriser des acides nucléiques in vitro ou in vivo. Différentes approches ont été décrites pour la production de virus recombinants, conduisant typiquement à la production de virus défectifs pour la réplication, comprenant un segment d'acide nucléique exogène codant un produit désiré. De tels virus ont été construits à partir
20 de rétrovirus (MLV, lentivirus, etc.), d'adénovirus (Ad5, Ad2, CAV, etc.), d'AAV, de virus de l'herpès, etc. Dans chacune de ces approches, un vecteur viral est construit comprenant les séquences nécessaires en cis à l'encapsidation d'un acide nucléique dans une particule virale et, éventuellement, des séquences régulatrices ou codantes complémentaires.

25

Dans le cas des rétrovirus, de nombreuses constructions ont été décrites, dans lesquelles tout ou partie des gènes gag, pol et/ou env sont délétées, et dans lesquelles un acide nucléique d'intérêt est introduit. Ce dernier peut être inséré en remplacement des séquences délétées, ou dans d'autres régions, comme par exemple dans un LTR. Des vecteurs viraux connus de l'homme du métier sont notamment MFG, pBABE, etc. Typiquement, un vecteur rétroviral
30 de l'invention comprend donc les régions terminales LTR, la séquence d'encapsidation, un gène de sélection et l'acide nucléique codant l'ARN actif, selon l'invention. Un tel vecteur

viral peut être construit à partir de différents types de rétrovirus, et utilisé pour produire des virus recombinants par introduction dans une lignée d'encapsidation exprimant les protéines virales GAG, POL et ENV. De telles lignées ont été décrites dans l'art antérieur (PsiCRIP, PA317, Gpenv, 293GP etc.).

5

Un objet particulier de l'invention réside dans un vecteur rétroviral déficient pour la réplication, comprenant une séquence LTR, une séquence d'encapsidation rétrovirale et au moins une cassette d'expression telle que définie précédemment.

- 10 Les vecteurs de l'invention peuvent comprendre une ou plusieurs cassettes d'expression d'ARN, identiques ou différentes. Ainsi, des vecteurs multi-cassettes peuvent être construits, dans lesquels des cassettes peuvent être insérées, par exemple en tandem. La possibilité de cloner plusieurs cassettes dans un même vecteur peut concerner une seule
- 15 la même cassette est recopiée plusieurs fois avant d'être insérée dans le vecteur ; dans le deuxième cas, les différentes cassettes sont placées les unes à côté des autres et insérées dans le vecteur, ou clonées séquentiellement, en des sites distincts du vecteur.

- Les vecteurs peuvent être construits par des techniques connues de biologie moléculaire,
- 20 notamment par clonage, ligation, amplification, etc.

- Comme indiqué précédemment, l'utilisation combinée d'un vecteur rétroviral et de séquences dérivées du gène VA d'un adénovirus, selon les méthodes décrites dans la présente demande, permet de fournir un système intégré, simple et prédictif de l'activité
- 25 d'ARN aléatoires, in vitro comme in vivo.

Un autre objet de l'invention concerne une composition comprenant un vecteur tel que défini précédemment. La composition peut être une composition pharmaceutique, comme il sera décrit plus en détail dans la suite du texte.

30

Un autre objet de l'invention concerne une composition comprenant une pluralité de vecteurs tels que définis précédemment. La composition peut être une banque, comme il sera décrit plus en détail dans la suite du texte.

5 Banque

Au sens de l'invention, le terme banque (ou librairie) désigne un produit ou une composition complexe comprenant une pluralité ou une multitude de composants ou de membres, qui peuvent être présents en mélange(s) ou dans des compartiments séparés. Les
10 banques selon l'invention comprennent typiquement une pluralité d'ARN actifs, ou de cassettes codant de tels ARN actifs, qui peuvent être clonées dans des vecteurs, notamment des plasmides, vecteurs viraux, virus, et/ou dans des cellules. Typiquement, bien que cela ne soit pas obligatoire, toutes les cassettes et/ou vecteurs d'une même banque ont sensiblement la même structure, ces composants différant les uns des autres par la nature (e.g., longueur,
15 origine, type, etc.) et/ou la structure (e.g., séquence) des ARN actifs codés. En outre, une banque comprend généralement plusieurs copies de chaque composant ou membre. La complexité de la banque peut varier dans une large mesure. Ainsi, une banque peut être composée de deux composants comprenant des cassettes d'expression d'ARN actifs distincts, de préférence de 10 au moins, encore plus préférentiellement de 20 au moins. Des
20 banques typiques comprennent plus de 100, 500 ou 1000 composants distincts, par exemple jusqu'à 10^9 ou plus encore. S'agissant de banques de cassettes d'expression d'ARN aléatoires, il est clair que la composition précise (e.g., la séquence) des composants de la banque est généralement et par principe non connue, au moins en partie. Les composants de la banque (e.g., cassettes, vecteurs, cellules, etc.) peuvent être présents sous des formes
25 diverses, comme sous forme liquide, gel, lyophilisée, etc. Ils peuvent être immobilisés sur un support ou en suspension, sous forme soluble. La banque comprend généralement un support physique contenant les différents membres de la banque, qui peuvent être en mélange(s), au moins partiel(s), ou séparés. Le support peut ainsi comprendre un ou plusieurs compartiments séparés physiquement, tels que des flasques, tubes, bouteilles,
30 plaques multi-puits, etc. La banque peut être conservée sous différentes formes, notamment en suspension liquide ou congelée, répliquée, etc., en tout ou en partie.

Les banques de l'invention comprennent typiquement une pluralité de vecteurs comprenant chacun une cassette d'expression d'un ARN aléatoire comme décrit précédemment, les vecteurs étant au moins en partie sous forme de mélange. De préférence, la banque
5 comprend au moins 50, 100 ou 200 vecteurs codant un ARN aléatoire distinct. Elle peut comprendre jusqu'à plusieurs milliards d'espèces moléculaires distinctes.

Les séquences aléatoires peuvent être toute molécule d'ADN ou d'ARN comprenant au moins un élément de séquence non connu, plus précisément toute molécule d'ADN ou
10 d'ARN dont une partie au moins de la séquence est aléatoire. De tels acides nucléiques aléatoires peuvent typiquement comprendre une région aléatoire bordée, à l'une ou aux deux extrémités, d'une région de séquence définie. La région aléatoire peut comprendre par exemple de 8 à 50 bases, et la ou les régions définies de 2 à 10 bases. L'acide nucléique aléatoire peut être un ARN simple-brin, produit par synthèse chimique, ou par amplification
15 ou mutagenèse à partir de toute matrice biologique, ou par expression d'un ADN correspondant. L'acide nucléique peut également être un ADN, notamment un ADN double-brin aléatoire codant un ARN aléatoire. Un tel ADN double-brin aléatoire peut être préparé à partir d'une population d'ADNs simple-brins aléatoires de synthèse ou obtenus par des techniques d'amplification et/ou de mutagenèse, par synthèse d'un second brin
20 complémentaire selon des techniques connues de l'homme du métier.

Un objet particulier de l'invention réside dans une banque d'acides nucléiques, caractérisée en ce qu'elle comprend une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression dérivée d'un gène VA1 d'un adénovirus
25 exprimant un ARN structural aléatoire distinct.

Un autre objet particulier concerne une banque d'acides nucléiques, caractérisée en ce qu'elle comprend une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression comprenant un ARN aléatoire distinct sous
30 contrôle d'un promoteur U6.

Un autre objet particulier concerne une banque d'acides nucléiques, caractérisée en ce qu'elle comprend une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression comprenant un ARN aléatoire distinct sous contrôle d'un promoteur tRNA.

5

Un objet particulier de l'invention réside dans une banque d'acides nucléiques, caractérisée en ce qu'elle comprend une pluralité d'ARN aléatoires distincts codés par des cassettes d'expression distinctes dérivées d'un gène VA1 d'un adénovirus.

- 10 Un objet particulier de l'invention réside dans une banque d'acides nucléiques, caractérisée en ce qu'elle comprend une pluralité de cassettes d'expression comprenant chacune une séquence codant un ARN structural aléatoire distinct placé sous le contrôle d'un promoteur transcrit par l'ARN polymérase III (notamment des cassettes d'expression dérivées d'un gène VA1 d'un adénovirus), chaque ARN structural aléatoire codé ayant la capacité in vitro
- 15 de lier une cible d'intérêt.

Un autre objet particulier de l'invention réside dans des banques d'acides nucléiques telles que définies ci-dessus, dans lesquelles les cassettes d'expression comprennent un promoteur VA inductible.

20

Un autre objet particulier de l'invention réside dans des banques d'acides nucléiques telles que définies ci-dessus, dans lesquelles les cassettes d'expression comprennent un second promoteur transcriptionnel, distinct et placé en amont du promoteur VA.

25 Sélection in Cellulo

Comme indiqué, l'invention concerne, de manière générale, des procédés de sélection in cellulo, à partir de banques d'acides nucléiques aléatoires, d'ARN actifs capables de conférer à une cellule un phénotype désiré.

30

Les procédés de l'invention comprennent, de manière générale :

- a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité de cassettes d'expression distinctes comprenant chacune une séquence nucléique codant un ARN aléatoire placée sous le contrôle d'un promoteur transcrit par l'ARN polymérase III,
- b) la mise en contact de ladite banque ou d'une partie de celle-ci avec une population
5 de cellules dans des conditions permettant le transfert d'acide nucléique dans lesdites cellules,
- c) la sélection d'une ou plusieurs cellules possédant le phénotype désiré, et
- d) l'identification de la ou des cassettes contenues dans la ou les cellules sélectionnées, ou des ARN actifs qu'elles expriment.

10

Dans un premier mode de réalisation, la banque mise en œuvre dans l'étape a) est une banque aléatoire générale, c'est-à-dire comprenant une pluralité de séquences totalement aléatoires. L'utilisation de ce type de banque est particulièrement intéressante pour la sélection d'ARN actifs capables de conférer un phénotype désiré à une cellule, sans
15 connaissance a priori de la cible biologique visée ou de la voie métabolique ciblée.

Dans un autre mode de réalisation, la banque mise en œuvre dans l'étape a) est une banque aléatoire restreinte, c'est-à-dire comprenant une pluralité de séquences dont le caractère aléatoire présente un certain niveau de restriction. Ainsi, la banque restreinte peut être une
20 banque dérivée de la séquence d'un gène cible donné, comprenant une multitude de séquences complémentaires d'une ou plusieurs régions de ce gène. La banque restreinte peut également être une banque aléatoire dans laquelle un ou plusieurs résidus, ou un ou plusieurs motifs de séquence sont imposés au sein de la région aléatoire. La banque restreinte peut également être une banque de mutants aléatoires d'une séquence cible
25 donnée ou une banque codant des ARN présélectionnés pour une propriété particulière. L'utilisation de banques aléatoires restreintes est particulièrement intéressante pour la sélection d'ARN actifs capables d'altérer une cible biologique déterminée ou une voie métabolique déterminée.

30 Ainsi, dans un mode de réalisation particulier, la banque mise en œuvre dans l'étape a) est une banque aléatoire restreinte codant des ARN aléatoires pré-sélectionnés pour une

propriété particulière, par exemple pour leur capacité à lier, in vitro, une cible d'intérêt (par exemple une protéine, un polypeptide, un peptide, un acide nucléique, une cellule, un lipide, etc.) ou pour leur affinité pour cette cible, pour une propriété propre, pour la présence d'un motif structural ou d'une séquence spécifique, etc. Dans ce contexte, un objet particulier de l'invention concerne un procédé de sélection, optimisation ou identification d'ARN actifs, comprenant :

1a) la préparation d'une banque d'acides nucléiques comprenant une pluralité de cassettes d'expression distinctes comprenant une séquence nucléique codant un ARN aléatoire placée sous le contrôle d'un promoteur transcrit par l'ARN polymérase III, les séquences d'ARN aléatoires codées ou les ARN complets contenant ces séquences aléatoires ayant été pré-sélectionnés in vitro pour leur capacité à lier (ou pour leur affinité pour) une cible d'intérêt,

1b) la mise en contact de cette banque ou d'une partie de cette banque, avec une population de cellules dans des conditions permettant le transfert d'acide nucléique dans lesdites cellules,

2) la sélection d'une ou plusieurs cellules possédant le phénotype désiré, et

3) l'identification de la ou des cassette(s) contenue(s) dans la ou lesdites cellules, ou de(s) l'ARN(s) actif(s) qu'elles expriment.

Les banques peuvent être produites par toute technique connue de l'homme du métier, notamment par synthèse, amplification, mutagenèse, etc., ou des combinaisons de ces méthodes. Il peut s'agir d'une banque d'ADN synthétiques ou produits par voie recombinante ou génétique, à partir de matrices artificielles ou synthétiques, comme des banque génomique, d'ARN de séquences obtenues par la méthode SELEX, ou par toute technique de mutagenèse ou d'évolution dirigée, etc.

Dans un mode de réalisation particulier, la banque d'ADN codant des ARN aléatoires est préparée par :

- synthèse d'une banque d'ADN simple-brin comprenant une région aléatoire encadrée par une ou deux régions de séquence définie,

- synthèse d'un deuxième brin au moyen d'une ADN polymérase et en présence d'une amorce complémentaire de la séquence définie du premier brin, ou d'une partie de celle-ci, pour produire une banque d'ADN double-brin comprenant une région aléatoire, et
- clonage de la banque d'ADN double-brin dans un vecteur, sous contrôle du promoteur choisi.

Ce procédé peut comprendre une étape supplémentaire d'expression et de sélection in vitro des ARN codés par la banque ayant la capacité d'interagir avec une cible biologique d'intérêt.

10

Dans un autre mode de réalisation particulier, la banque d'ADN codant des ARN aléatoires est préparée à partir d'une collection de séquences ARN aléatoires par :

- transcription inverse pour produire une banque d'ADN simple-brin comprenant une région aléatoire encadrée par une ou deux régions de séquence définie,
- synthèse d'un deuxième brin au moyen d'une ADN polymérase et en présence d'une amorce complémentaire de la séquence définie du premier brin, ou d'une partie de celle-ci, pour produire une banque d'ADN double-brin comprenant une région aléatoire, et
- clonage de la banque d'ADN double-brin dans un vecteur, sous contrôle du promoteur choisi.

20

Dans un mode de mise en œuvre préféré, le vecteur est un vecteur viral, notamment rétroviral. Dans ce cas, le procédé comprend avantageusement en outre une étape de transfection dudit vecteur dans une lignée cellulaire d'encapsulation, pour produire une banque de virus, notamment de rétrovirus recombinants.

25

Dans un autre mode de réalisation particulier, la banque d'ADN code des ARN aléatoires pré-sélectionnés in vitro (voir Figure 15B). Dans ce mode de réalisation, le procédé peut comprendre les étapes suivantes:

- Etape A : la synthèse in vitro d'une banque de cassettes d'expression (ADN db) utilisant le système polymérase III. A cet effet, on peut par exemple introduire la séquence active aléatoire à l'intérieur d'une cassette d'expression telle que définie précédemment. Les

30

cassettes peuvent être synthétisées par exemple à partir d'oligonucléotides ADN sb, grâce à des réactions d'élongation d'amorce et/ou d'amplification par PCR;

- Etape B : l'expression in vitro de la banque de cassettes ou d'une partie de celle-ci, produisant in vitro une banque d'ARN aléatoires ;

5 - Etape C : la sélection in vitro des ARN pour leur capacité de liaison ou pour leur affinité pour une cible donnée; et

- Etape D : la production d'une banque de cassette d'expression restreinte comprenant des cassettes d'expression d'ARN ainsi sélectionnés.

10 L'expression in vitro de la banque de cassettes peut être réalisée en deux phases, une phase de production de cassettes de transcription et une phase de transcription proprement dite. Pour cela, une banque de cassettes de transcription in vitro peut être synthétisée à partir des banques de cassettes d'expression, totalement in vitro, en utilisant des oligonucléotides ADN sb et par réaction PCR. La matrice ADN db étant constituée par la banque de cassettes
15 d'expression, l'oligonucléotide en 5' utilisé dans la réaction PCR permet d'apporter un promoteur adapté à la transcription in vitro (ex : SP6, T7, T3,...). La production de la banque d'ARN aléatoires peut ensuite être effectuée par transcription in vitro à partir de la banque de cassettes de transcription, en utilisant une ARN polymérase adaptée (protéine purifiée ou préparation qui contient l'activité requise : SP6, T7, T3, ...).

20

La production de la banque de cassette d'expression restreinte peut être réalisée de différentes façons. D'une manière pratique, les ARN actifs sélectionnés (ou une partie de ceux-ci) sont utilisés pour générer les cassettes de transcription correspondantes, par exemple par réaction de RT-PCR. Les nouvelles banques de cassettes ainsi obtenues sont
25 utilisées soit pour effectuer de nouvelles itérations du procédé (retour à l'étape B), soit pour des tests in cellulo. Les tests in cellulo peuvent concerner quelques cassettes identifiées ou de façon plus globale cette banque de cassette d'expression restreinte peut être utilisée comme matériel de départ de la sélection in cellulo (étape 1b).

30 Un objet particulier de l'invention réside également dans un procédé de sélection d'ARN actifs, comprenant (i) la synthèse in vitro d'une banque de cassettes d'expression (ADN db)

codant pour des ARN aléatoires sous contrôle d'un promoteur polymérase III, (ii) l'expression in vitro de la banque de cassettes ou d'une partie de celle-ci, produisant in vitro une banque d'ARN aléatoires et (iii) la sélection in vitro des ARN pour leur capacité de liaison ou pour leur affinité pour une cible donnée. Dans une étape supplémentaire
5 facultative, le procédé comprend la production d'une banque de cassettes d'expression restreinte comprenant des cassettes d'expression d'ARN ainsi sélectionnés.

Lors de l'étape b) du procédé de l'invention, la banque (ou une partie de celle-ci) est mise en contact avec une population de cellules. Les cellules utilisées peuvent être de nature et
10 d'origine variées, et choisies en fonction des propriétés recherchées pour l'ARN actif. Ainsi, le procédé de l'invention peut être mis en œuvre notamment avec une population de cellules comprenant des cellules animales (par exemple de mammifère), d'oiseaux, de poissons, de batraciens, de plante, d'insecte, de levure ou bactériennes. Il s'agit de préférence de cellules de mammifères, notamment humaines ou d'animaux (rongeurs, bovins, équins, singes, etc.).
15 Les cellules peuvent être des cultures primaires ou des lignées. Il peut s'agir de cellules pluripotentes embryonnaires ou somatiques, différenciées ou non, prolifératives ou quiescentes, etc. On peut citer par exemple des cellules souches, des fibroblastes, des hépatocytes, des cellules épithéliales, musculaires, rénales, nerveuses, cardiaques, ou appartenant au lignage hématopoïétique (lymphocytes B, T, NK, mastocytes, cellules
20 dendritiques, macrophages résidants ou circulants, etc), etc. Les cellules utilisées peuvent par ailleurs être modifiées ou traitées préalablement, par exemple pour contenir un système de gène rapporteur, un marqueur, etc., ou pour présenter un phénotype pathologique que l'on souhaite corriger.

25 Le procédé de sélection est typiquement réalisé in vitro, dans tout type de support adapté, tel que fiole, flasque, plaque multi-puits, etc. A cet effet, afin de pouvoir mesurer ou observer, sur chaque cellule, l'effet d'un nombre restreint d'ARN aléatoires de la banque, la banque est préférentiellement mise en contact avec la population de cellules dans des conditions permettant le transfert d'un nombre restreint de cassettes par cellule. En effet, la banque
30 étant typiquement composée d'un mélange de composants distincts, on préfère que chaque cellule de la population soit modifiée par un nombre restreint de ces composants pour mieux

en apprécier les propriétés. De ce fait, il n'est pas nécessaire que les constituants de la banque soient séparés les uns des autres, ou que la population de cellules soit répartie en supports à plusieurs compartiments, ce qui constitue un avantage important de l'invention.

A titre indicatif, lorsque la banque est une banque de virus, on préfère incuber les cellules à une MOI faible, typiquement inférieure à 5, de préférence inférieure à 3, 2 ou 1.

La mise en contact peut être réalisée en présence d'agents facilitant la transfection, comme des polymères, lipides cationiques, peptides, etc. Lorsque la banque comprend des virus, notamment des rétrovirus, de tels agents ne sont généralement pas nécessaires, compte tenu de l'efficacité d'infection.

Les cellules peuvent être cultivées ou conservées un certain temps après la mise en contact, avant de réaliser l'étape c). Cette durée peut être ajustée par l'homme du métier selon le phénotype recherché, le type de vecteur, la quantité de cellules, etc. Par ailleurs, consécutivement à la mise en contact, il est possible de réaliser une étape de sélection des cellules dans lesquelles une ou plusieurs cassettes ont effectivement été transférée. Cette sélection peut être effectuée par tout moyen connu de l'homme du métier, notamment en utilisant un gène marqueur inséré sur le vecteur. En outre, les cellules peuvent également être soumises à des traitement ou conditions particulières, notamment pour révéler le phénotype d'intérêt (e.g., ajout d'un substrat, d'un réactif, lyse des cellules, etc.)

Le phénotype d'intérêt peut être toute activité, propriété, morphologie, etc. Il peut s'agir de l'expression d'un gène endogène ou exogène, d'un marqueur, de l'expression d'une protéine de surface, d'une propriété de migration, différenciation, croissance, résistance, etc. Dans un mode de mise en œuvre particulier, le phénotype désiré est choisi parmi une capacité ou une incapacité de croissance, d'apoptose, de différenciation, de migration, de résistance à un agent toxique, de résistance à un agent infectieux, ou d'action métabolique (e.g., la cellule est devenue capable de modifier son environnement métabolique). Dans un autre mode de réalisation, le phénotype désiré est l'activité d'une cible biologique déterminée ou d'une voie métabolique déterminée. On peut citer notamment l'expression ou

l'activité d'une protéine, par exemple d'une enzyme (e.g. kinase, protéase, etc.), d'un facteur de transcription, etc.

Dans un premier mode de réalisation, la population de cellules comprend des cellules
5 infectées par un virus et le phénotype désiré est la résistance audit virus. Le virus peut être tout virus connu, comme un virus de l'hépatite (B, C, delta...), de la grippe, le VIH, les différents herpès, les virus des papillomes, etc.

Dans un autre mode de réalisation, la population de cellules comprend des cellules
10 tumorales et le phénotype désiré est la perte de tumorigénicité.

Dans un autre mode de réalisation, la population de cellules comprend des cellules souches embryonnaires non différenciées et le phénotype désiré est le contrôle de leur différenciation.

15 Dans un autre mode de réalisation, la population de cellules comprend des cellules capables d'agir sur un processus métabolique naturel (par exemple : coagulation sanguine, régulation du taux de glucose, de lipides, cholestérol...) et le phénotype désiré est le contrôle de ce processus métabolique.

20 Selon une autre variante, la population de cellules comprend des cellules bactériennes et le phénotype désiré est la sensibilité à un agent toxique.

Dans un autre mode de réalisation, la population de cellules comprend des cellules
25 exprimant une cible biologique déterminée (e.g. une protéine, un variant d'une protéine, un acide nucléique, un lipide, un récepteur, etc.) et le phénotype désiré est la modification de l'activité (y compris de l'expression) de cette cible biologique.

Les cellules exprimant le phénotype désiré peuvent être sélectionnées par l'homme du
30 métier par toute technique classique de biologie (modification morphologique, survie, expression d'un marqueur, tri cellulaire, etc.). En outre, lorsque la cassette est inductible,

l'activité de l'ARN peut être validée directement in cellulo, en comparant les états induits et réprimés (Figure 15A). Pour cela, les cellules présentant le phénotype désiré sont sélectionnées, éventuellement amplifiées, de préférence de manière individuelle, et leur phénotype est analysé en parallèle en conditions d'induction et de répression de l'expression de la cassette. Cette étape supplémentaire permet d'identifier les ARN dont l'activité est
5 directement impliquée dans le phénotype recherché.

L'étape d) comprend l'identification de la ou des cassettes contenues dans la ou les cellules sélectionnées, ou des ARN actifs qu'elles expriment. Ces cassettes ou ARN sont
10 responsables du phénotype produit et peuvent donc être utilisés pour toute application visant à reproduire ce phénotype. La cassette, ou l'ARN peuvent être extraits des cellules et isolés par des méthodes classiques de biologie moléculaire (lyse des cellules, amplification ou hybridation, etc.). Dans un mode de mise en œuvre préféré, la séquence de la ou des cassettes est déterminée, pour permettre de produire par voie synthétique ou recombinante
15 le produit correspondant. Bien entendu, les propriétés de la cassette ou de l'ARN peuvent être confirmées dans tout système ou modèle biologique approprié.

Préférentiellement, lorsque la banque utilisée dans l'étape a) est complexe (i.e., comprend un nombre élevé de composants distincts, par exemple supérieur à 100), on préfère répéter
20 les étapes b) à d) du procédé afin de sélectionner des agents les plus actifs. Dans ce cas, l'ADN des cassettes d'expression des cellules sélectionnées est amplifié pour produire une banque restreinte, et les étapes b)-d) du procédé sont répétées une fois au moins avec ladite banque restreinte. La réalisation de plusieurs cycles présente plusieurs avantages : tout d'abord, elle permet de partir de banques très complexes, utilisées sous forme de mélange.
25 De plus, elle permet d'augmenter progressivement l'efficacité des ARN actifs. Par ailleurs, elle peut permettre de sélectionner des ARN actifs présentant un profil déterminé, en sélectionnant les cassettes sur des cellules ou dans des conditions distinctes selon les cycles. Ainsi, la répétition de cycles peut permettre de contrôler la spécificité d'un ARN actif ou au contraire de vérifier son efficacité sur plusieurs cibles ou plusieurs types cellulaires.

Un objet particulier de l'invention réside dans un procédé de sélection d'ARN actifs capables de conférer à une cellule un phénotype désiré, comprenant :

- a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité de vecteurs comprenant des cassettes d'expression distinctes comprenant chacune une séquence
5 nucléique codant un ARN aléatoire placée sous le contrôle d'un promoteur transcrit par l'ARN polymérase III,
- b) la mise en contact de ladite banque, ou d'une partie de celle-ci, avec une population de cellules dans des conditions permettant le transfert d'acide nucléique dans lesdites cellules,
- 10 c) la sélection de cellules possédant le phénotype désiré,
- d) l'extraction ou l'amplification de la séquence de cassettes contenues dans lesdites cellules,
- e) le clonage des séquences obtenues en d) dans un vecteur pour générer une banque restreinte et,
- 15 f) la répétition une fois au moins des étapes b) à d) avec ladite banque restreinte.

Dans un mode préféré, la banque d'acide nucléiques est une banque codant des ARN aléatoires pré-sélectionnés in vitro, et/ou le vecteur est un virus recombinant, plus préférentiellement un rétrovirus recombinant.

- 20 De manière particulièrement préférée, le promoteur transcrit par l'ARN polymérase III est un promoteur dérivé de la séquence d'un gène VA d'un adénovirus.

Dans un mode spécifique, la population de cellules comprend des cellules de mammifère.

Un mode de réalisation plus particulier de l'invention comprend :

- 25 a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression dérivée d'un gène VA d'un adénovirus exprimant un ARN structural aléatoire distinct,
- b) la mise en contact de ladite banque ou d'une partie de celle-ci avec une population
30 de cellules de mammifères dans des conditions permettant l'infection desdites cellules par lesdits rétrovirus recombinants,

- c) la sélection des cellules possédant le phénotype désiré,
- d) l'extraction ou l'amplification de la séquence de cassettes contenues dans lesdites cellules,
- e) le clonage des séquences obtenues en d) dans un vecteur pour générer une banque
- 5 restreinte et,
- f) la répétition une fois au moins des étapes b) à d) avec ladite banque restreinte.

Un autre objet particulier de l'invention réside dans un procédé de sélection d'ARN actifs sur une cible biologique déterminée, comprenant :

- 10 a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité de vecteurs comprenant des cassettes d'expression distinctes comprenant chacune une séquence nucléique codant un ARN contraint (ou prédéfini) pour agir sur ladite cible déterminée, placée sous le contrôle d'un promoteur transcrit par l'ARN polymérase III,
- b) la mise en contact de ladite banque, ou d'une partie de celle-ci, avec une population
- 15 de cellules exprimant ou contenant la cible biologique, dans des conditions permettant le transfert d'acide nucléique dans lesdites cellules,
- c) la sélection de cellules possédant le phénotype désiré,
- d) l'extraction ou l'amplification de la séquence de cassettes contenues dans lesdites cellules et, de manière facultative,
- 20 e) le clonage des séquences obtenues en d) dans un vecteur pour générer une banque restreinte et,
- f) la répétition une fois au moins des étapes b) à d) avec ladite banque restreinte.

- La séquence des ARN codés peut-être dérivée de la séquence de la cible biologique
- 25 (notamment dans le cas des antisens, ARNi (siARN, miARN ou leurs précurseurs), ribozymes) ou bien être pré-sélectionnée pour interagir de manière structurale avec la cible biologique (notamment dans le cas des aptamères).

Applications pour la recherche en biotechnologie

Les ARN actifs identifiés, les séquences actives identifiées ou les cassettes d'expression de ces ARN actifs peuvent être utilisés comme des outils moléculaires capables d'agir dans la cellule pour interférer (inhibition, activation) avec une activité biologique ou l'expression d'un phénotype déterminé (« target identification »). A ce titre, ils constituent des produits utiles pour étudier un processus cellulaire et identifier de nouvelles cibles ou pour explorer la fonction d'un gène dans le cas où la cible est connue (« target validation »). Leur action dans une cellule peut permettre de modifier la cellule de façon à ce que ladite cellule soit dotée de propriétés nouvelles. La cellule ainsi modifiée peut alors être considérée comme un nouvel outil en biotechnologie utilisable à des fins de recherche ou pour des applications thérapeutiques.

Applications pharmaceutiques.

Les ARN actifs identifiés et, plus généralement, les cassettes d'expression identifiées, sont utilisables directement comme produits pharmaceutiques. Dans le contexte de l'invention, le terme « pharmaceutique » inclut toute utilisation dans le domaine médical, thérapeutique, préventif ou curatif, vétérinaire, agronomique, diagnostique, cosmétique, etc.

Ainsi, un aspect de l'invention concerne une composition pharmaceutique comprenant une cassette d'expression, un vecteur ou une cellule tels que définis ci-avant, et un véhicule ou excipient acceptable sur le plan pharmaceutique.

Un autre aspect de l'invention concerne une composition pharmaceutique, caractérisée en ce qu'elle comprend un ARN actif, ledit ARN actif comprenant une séquence active insérée dans un ARN VA modifié, ledit ARN VA modifié comprenant éventuellement une hélice terminale altérée et/ou une séquence conférant un caractère inductible.

L'invention concerne encore des procédés de production de compositions pharmaceutiques, comprenant (i) le criblage d'une banque d'ARN aléatoires comme décrit ci-avant,

permettant d'obtenir une cassette d'expression d'un ARN actif et (ii) le conditionnement de la cassette d'expression ou de la séquence ARN active dans tout excipient ou véhicule acceptable sur le plan pharmaceutique.

- 5 Dans un mode de mise en œuvre particulier, l'invention concerne un procédé de production d'une composition pharmaceutique pour le traitement d'une infection par un agent pathogène chez un patient humain, comprenant (i) le criblage d'une banque d'ARN aléatoires comme décrit ci-avant, la population de cellules utilisées étant infectée par l'agent pathogène et les ARN sélectionnés pour leur capacité à réduire ou bloquer le cycle infectieux, permettant d'obtenir une cassette d'expression d'un ARN actif et (ii) le
10 conditionnement de la cassette d'expression ou de la séquence ARN active dans tout excipient ou véhicule acceptable sur le plan pharmaceutique.

- L'invention concerne aussi l'utilisation d'un ARN actif, d'une cassette d'expression, d'un
15 vecteur ou d'une cellule recombinante tels que définis ci-avant, pour la préparation d'un médicament destiné à la mise en œuvre d'une méthode de traitement thérapeutique du corps humain. Selon les propriétés de l'ARN actif, le médicament peut être utilisé pour le traitement de cancers, infections, maladies neurodégénératives, etc.

- 20 L'invention concerne encore une méthode de traitement d'un patient, comprenant l'administration d'une quantité efficace d'un ARN actif, d'une séquence active, d'une cassette d'expression, d'un vecteur ou d'une cellule recombinante tels que définis ci-avant à un patient. L'administration peut être réalisée par différentes voies, notamment par voie iv, ip, im, sc, locale ou générale, notamment intra-tumorale ou systémique.

25

D'autres aspects et avantages de la présente invention apparaîtront à la lecture des exemples qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

30

EXEMPLES

oA. CONSTRUCTION DES CASSETTES D'EXPRESSION

5 A-1. CASSETTES DERIVEES DU GENE VA1 DES ADENOVIRUS.

Cet exemple décrit la construction de cassettes permettant l'expression et la diffusion intracellulaire (en cellules de mammifères) de motifs d'ARN actifs (ARN structuraux = aptamères, antisens, ribozymes, ARNi (siARN ou miARN) ou motifs actifs).

- 10 La base de ces cassettes est le gène viral VA1 RNA de l'adénovirus de type 2. Ce gène est efficacement transcrit par l'ARN polymérase III cellulaire. L'ARN produit est très structuré, sa taille est de 160 bases. La localisation cellulaire de cet ARN est cytoplasmique.

- Afin d'éliminer la partie physiologiquement active de l'ARN VA1, la boucle IV (qui interagit avec la protéine kinase p68) a été supprimée dans cette construction et le site de restriction EcoRV a été inséré : construction VAAIV (Barcellini et al., 1998) et (Gwizdek et al., 2001).

- L'ARN VAAIV a une taille de 134 bases, il est riche en structures secondaires (figure 4) et a une localisation cytoplasmique (Barcellini et al., 1998) (Gwizdek et al., 2001). Sa séquence est représentée par la séquence SEQ ID NO : 2.

A-1-a. Modification de la cassette VAAIV : Cassette VAAIV (cytoplasmique) et nVAAIV (nucléaire).

25

Cassette VAAIVSrf :

Cet exemple décrit la construction de cassettes permettant l'expression des motifs actifs intégrés dans l'ARN VAAIV, et dont la localisation soit principalement cytoplasmique.

- 30 Dans le domaine central du gène VA1, au niveau de la délétion de la boucle IV, le site de clonage SrfI, plus adapté, a été introduit en remplacement du site EcoRV.

La cassette VAAIVSrf a été générée en insérant le site SrfI sous la forme d'un octanucléotide ADN double brin (5'GCCCCGGGC3') à l'intérieur du site de restriction EcoRV (position 90-96 dans le VAAIV). L'ARN transcrit possède 142 bases (figure 5) et une localisation cytoplasmique (figure 2C), SEQ ID NO : 3. Cet ARN permet l'expression optimisée de séquences ARN actives (figure 2B).

Cassette nVAAIVSrf :

Cet exemple décrit la construction de cassettes permettant d'exprimer des séquences d'ARN actives intégrées dans l'ARN VAAIVSrf, et dont la localisation reste nucléaire.

Dans l'ARN VA1 natif, la structure en double hélice contenant les extrémités 5' et 3' est la séquence responsable du transport de cet ARN du noyau vers le cytoplasme. Cette structure est appelée l'hélice terminale (Gwizdek et al., 2001). Elle a été cartographiée : bases 1 à 20 et bases 136 à 155. Si la structure en double hélice est perturbée, l'ARN est retenu dans le cytoplasme. La partie 3' du gène VAAIV a ainsi été modifiée afin de créer une rupture dans la double hélice terminale de l'ARN selon l'approche proposée dans Gwizdek et al., 2001. Pour cela, les séquences 5' du gène VAAIV ont été modifiées à partir du nucléotide 93 et remplacées par la séquence ADN double brin suivante (SEQ ID NO : 5) :

5'GCCCCGGGCATCCAGGTGTGCGACGTCAATAAACGGGGGAGCGCCCTTTT3'.

On peut noter que le site de restriction SrfI, en italique et souligné, est directement intégré dans cette séquence (figure 6A). L'ARN transcrit dont les caractéristiques sont présentées dans la figure 6B possède 141 bases et une structure secondaire proche de celle de l'ARN VA1, sa localisation est nucléaire (figure 6C), SEQ ID NO : 4.

A-1-b. Inductibilité de la cassette VAAIV : Cassettes VAI (cytoplasmique) et nVAI (nucléaire).

Cassettes VAI :

5

Cet exemple décrit la construction de cassettes permettant d'exprimer des motifs actifs intégrés dans une cassette d'expression inductible par la tétracycline.

Premièrement, le gène VAAIV a été modifié afin d'intégrer une séquence opératrice tetO1 entre les Boite A et B du gène VA et/ou en amont du site d'initiation de la transcription de
10 ce gène (figure 7). La séquence tetO1 remplace donc les séquences natives du gène VA soit entre les boîte A et B (position 24 à 59 dans le gène VAIo) soit en amont des séquences transcrites (position - 29 à - 50 dans le gènes OVAi), soit au niveau des deux positions (extragénique et intragénique : gène OVAio). Deuxièmement, dans le domaine central de ce nouveau gène VA se trouve le site de clonage SrfI. Et troisièmement, dans le but de recréer
15 une structure secondaire proche de celle de l'ARN VA natif, des séquences partiellement complémentaires à la séquence tetO1 remplacent les séquences centrales du gène (VAio et OVAio). Afin d'obtenir un ARN VAI cytoplasmique tout est fait pour conserver l'hélice terminale responsable du transport de l'ARN vers le cytoplasme. La séquence complète des gènes OVAi (SEQ ID NO : 16), OVAio (SEQ ID NO : 17) et VAIo (SEQ ID NO : 15) est
20 indiquée dans la figure 7A.

Les cassettes VAI sont obtenues de manière synthétique en utilisant 4 oligonucléotides d'ADN simple brin. Les deux premiers oligonucléotides d'ADN simple brin VAI *up* et VAI *down* (SEQ ID NOs : 18 et 19, figure 7B) sont utilisés pour générer la totalité des séquences
25 transcrites du néogène VAI ainsi qu'une partie des séquences adjacentes en 5' et en 3'. Grâce à une séquence de 20 bases complémentaires ces oligonucléotides sont hybridés puis utilisés pour générer un ADN double brin en présence de l'ADN polymérase fragment de Kleenow. Les deux oligonucléotides externes VAI PvuII 5' (SEQ ID NO : 20) et VAI PvuII 3' (SEQ ID NO : 22) (figure 7B) sont utilisés pour ajouter les séquences amont et aval du
30 gène VA grâce à une réaction de PCR. Les deux sites de restriction PvuII placés aux deux extrémités 5' et 3' du gène VAIo permettent de le cloner dans un vecteur ADN quelconque.

De même, les deux oligonucléotides VAI02PvuII 5' et VAIpvuII 3' sont utilisés pour ajouter la séquence tetO1 en amont du gène VAAIVSrf ou du gène VaiO afin de générer respectivement les gènes OVAi et OVAiO.

- 5 L'ARN transcrit VAI0 possède 142 bases, sa structure secondaire est présentée en figure 8.

Cassette nVAi :

- 10 Cet exemple décrit la construction de cassettes permettant d'exprimer des motifs actifs intégrés dans une cassette d'expression inductible par la tétracycline. La localisation de ce motif doit être principalement nucléaire. Le principe de la rétention de cet ARN dans le noyau est le même que pour la cassette d'expression nVAAIVSrf (voir ci-dessus). La partie 3' des différents gènes VAI est donc modifiée de la même façon que pour le gène nVAAIVSrf afin de créer une rupture dans la double hélice terminale de l'ARN.

15

A-2. CASSETTES DERIVEES DU GENE HUMAIN hU6

- 20 La séquence promotrice du gène U6 humain est en position extragénique. L'ARN généré correspond donc aux séquences que l'on a choisi de cloner en aval du promoteur (à partir du point +1 de transcription). Ce type de promoteur permet donc d'exprimer des ARN entièrement synthétiques par opposition au système VA qui impose de conserver des séquences promotrices positionnées au sein des séquences transcrites.

- 25 Cet exemple décrit plusieurs modifications du gène U6+1 (décrit initialement par (Bertrand et al., 1997)). Ces modifications permettent de faciliter le clonage des séquences actives, de maîtriser les séquences transcrites dès le point +1 de transcription, de stabiliser les séquences actives et de maîtriser la localisation intracellulaire de ces séquences.

Les séquences promotrices utilisées dans les cassettes vont de 1 à 266. Au niveau du site +1 de transcription (nu.266), un site de clonage à bouts collants a été placé : Sall (5'GTCGAC3') (Bertrand et al., 1997).

5 A-2-a. Cassettes hélice U6 (U6h9, U6h20 et nU6h20)

Les ARN transcrits par ces cassettes ont une localisation nucléaire ou cytoplasmique, en fonction des constructions. L'objectif de ces cassettes est l'expression intracellulaire de motifs d'ARN structuraux.

10

Une séquence servant de base structurale pour l'ARN à insérer est placée en aval du promoteur U6. Elle contient : une séquence capable de générer une courte hélice ARN (plus ou moins stable et plus ou moins longue), un site de clonage bord franc (SrfI : 5'GCCCCGGGC3') ainsi que le signal d'arrêt de transcription TTTT.

15

Dans les cellules de mammifères, l'accumulation d'ARN double brin (d'une longueur supérieure à 40 nucléotides environ) déclenche la production d'interféron et la mort des cellules par apoptose. Cependant il est nécessaire que la navette ARN qui supporte le motif actif possède une structure propre et stable en forme d'hélice. Cet exemple décrit la construction de trois types de navette dérivées de U6 : lorsque le motif actif est lui même structuré sous la forme d'une hélice, nous utilisons une cassette possédant un court motif navette (hélice 9 : h9U6), lorsque le motif actif n'est pas structuré en hélice ou très court, nous utilisons une cassette permettant l'expression d'une navette formé par l'appariement d'une vingtaine de nucléotides (hélice 20 h20U6). h20U6 permet également de modifier la tige terminale de la navette motif d'export et de générer une navette à localisation nucléaire (nh20U6).

20

25

Différents fragments d'ADN double brin sont insérés à l'intérieur du site de restriction Sall grâce à des extrémités Sall compatibles.

30

h9U6 : (SEQ ID NO : 6)

5' TCGAGCCCCGGGCTCGACTTTTTTC 3'

3' CGGGCCCGAGCTGAAAAAGAGCT 5'

Séquence transcrite du gène h9U6 : (SEQ ID NO : 7)

5 GTCGAGCCCCGGGCTCGACTTTTT

h20U6: (SEQ ID NO : 8)

5' TCGAGGATATCGACTGCCCCGGGCAGTCGATATCCTCGACTTTTTTC 3'

3' CCTATAGCTGACGGGCCCCGTCAGCTATAGGAGCTGAAAAAGAGCT 5'

10 Séquence transcrite du gène h9U20 : (SEQ ID NO : 9)

GTCGAGGATATCGACTGCCCCGGGCAGTCGATATCCTCGACTTTTT

nh20U6 :

(SEQ ID NO : 10)

15 5' TCGAGGATATCGACTGCCCCGGGCAGAGATAAGGTCGACTTTTTTC 3'

3' CCTATAGCTGACGGGCCCCGTCTCTATTCCAGCTGAAAAAGAGCT 5'

L'ARN correspondant à une structure en hélice sur 18 bases, avec une interruption d'hélice sur 6 bases. Sa séquence est indiquée ci-dessous (SEQ ID NO : 11)

GTCGAGGATATCGACTGCCCCGGGCAGAGATAAGGTCGACTTTTT

20

La structure secondaire des hélice U6 est présentée dans la figure 9.

A-2-b. Cassette U6 Srf

25 L'objectif de cette cassette est d'exprimer des ARN dont la séquence est choisie dès le point +1 de transcription.

Dans cette cassette, le site de restriction à extrémités cohésives Sal I est remplacé par le site de restriction à bords francs SrfI.

30 Pour cela, le promoteur U6 est modifié par réaction PCR avec pour matrice la cassette U6 décrite précédemment. Dans cette réaction, l'amorce 5' est une séquence plasmidique située

en amont du promoteur et l'amorce 3' permet de modifier les séquences à proximité du site +1 de transcription :

5'GTGGGCCCATGGGTGCCCCGGGCTTTCGTCCTTTCCACAAG3' (SEQ ID NO :12).

Dans cet oligo, la séquence Srf I GCCCCGGGC remplace la séquence CACCGTCG présente dans le gène d'origine (Bertrand et al., 1997) (5' CACCGTCG 3' de la cassette d'origine a été changée par le site SrfI 5' GCCCCGGGC 3' (les séquences soulignées représentent le début des séquences transcrites).

La taille de l'ARN transcrit est fonction de la séquence clonée en aval du promoteur, à l'intérieur du site SrfI. De même, la localisation cellulaire de l'ARN transcrit est fonction de la séquence de ce dernier.

A-2-c. Cassette U6Tt

L'Objectif de cette cassette est d'exprimer des séquences d'ARN actif (aptamère, antisens, ribozyme, ARNi : siARN ou miARN) qui ne soient pas dégradés par les RNAses cellulaires.

En aval du promoteur U6, à l'intérieur du site de restriction SalI, est clonée une séquence appelée tige terminale. Elle génère une structure en épingle à cheveux à l'extrémité 3' de l'ARN, évitant ainsi la dégradation par les RNAses. En amont de cette tige terminale se trouve la séquence GTCGAC qui reconstitue un site de restriction SalI, afin de représenter le site de clonage des séquences antisens. En aval de la tige terminale se trouve le signal d'arrêt de transcription de la polymérase III : TTTTT.

Séquence de la tige terminale (SEQ ID NO : 13)

5' GCGGACTTCGGTCCGCTTTTTT 3'

Les séquences soulignées forment l'hélice ARN

La séquence transcrite dans cette cassette est (SEQ ID NO : 14)

5' G//TCGACCCATGCTAGAGCGGACTTCGGTCCGCTTTTT

// représente le site d'insertion SalI pour les séquences actives.

Les séquences en gras représentent la tige terminale.

A-3. CASSETTES HYBRIDES U6/VA

5 Cet exemple illustre la construction d'un gène hybride utilisable en particulier dans les lignées murines. En effet, le gène VA1, ainsi que les cassettes d'expression dérivées de ce gène, ne s'exprime pas dans les lignées murines. Pour pallier ce manque d'expression, le promoteur du gène U6 murin est utilisé pour transcrire les cassettes d'expression dérivées du gène VA1.

10

La construction mU6/VAiO a été effectuée en deux étapes : la première a consisté à insérer le promoteur U6 du gène murin (mU6) dans le vecteur rétroviral pBabe et la deuxième a consisté à insérer le gène VAiO en aval du promoteur mU6.

15 Étape N°1 : Clonage de mU6 dans pBabe (Figure 14)

Le gène U6 murin a été copié par réaction de polymérisation en chaîne grâce aux amorces mU6 amont et mU6 aval. En 3' de l'amorce mU6 aval, 4bases supplémentaires ont été ajoutées afin d'intégrer un site de restriction. Le site de restriction choisi est le site Pme I de séquence GTTTAAAC. Ce site de restriction présente deux avantages : il permet i) de
20 conserver intactes les dernières bases du promoteur U6 murin (GTTT) et ii) d'intégrer un site de coupure bord franc utilisable dans l'étape N°2 (insertion du gène VAiO). La réaction de polymérisation en chaîne a été effectuée avec pour matrice de l'ADN génomique extrait de cellules murines. Le produit de réaction a été purifié puis inséré dans le vecteur rétroviral pBabe au niveau du site de restriction NheI (figure 10) pour générer le plasmide pBabe
25 mU6.

Étape N°2 : Clonage de VAiO dans pBabe mU6 (Figure 14)

Le gène VAiO, obtenu par réaction de polymérisation en chaîne, a été cloné dans le plasmide pBabe mU6 au niveau du site de restriction PmeI précédemment introduit. Les
30 oligonucléotides utilisés pour l'obtention du gène VAiO sont VAiO 5' et VAiO End NheI,

un site de restriction Nhe I étant introduit en 3' de l'amorce VaiO End NheI. La matrice d'ADN utilisée était le plasmide pBabe VaiO.

5 B. CLONAGE DES CASSETTES D'EXPRESSION DANS UN VECTEUR

B-1. VECTEUR RETROVIRAL pBabe

Le vecteur rétroviral choisi est le vecteur pBabe (Morgenstern and Land, 1990). Le site d'insertion choisi est le site NheI qui est localisé dans le LTR 3' (figure 10). Ce site d'insertion présente plusieurs avantages :

- Au cours du cycle viral, le LTR 3' est celui qui est copié pour générer les deux nouveaux LTR du provirus intégré. Il est donc responsable de l'activité du promoteur viral en situation d'intégration dans l'ADN cellulaire. Le fait d'intégrer une séquence exogène à l'intérieur de ce LTR ne perturbe pas la production des virus recombinants par les cellules d'encapsidation, mais permet d'inactiver le promoteur viral en situation intégrée dans les cellules transduites. L'intégration de telles séquences dans un génome cellulaire n'entraîne donc pas l'activation des gènes situés en aval de l'insertion. Ce type de construction est particulièrement avantageux dans l'optique d'une utilisation pour des projets de thérapie génique ou cellulaire.
- Une des conséquences de la duplication du LTR 3' est aussi la duplication de la séquence insérée dans ce LTR. Il y a donc deux copies des cassettes d'expression par intégration.

Les cassettes construites sont ainsi transférées dans le vecteur pBabe au niveau du site d'insertion NheI (à l'intérieur du LTR 3').

B-2. AUTRES VECTEURS

Tout autre vecteur ou moyen permettant de faire pénétrer les cassettes d'expression dans la cellule est également envisagé : transfection, transduction ou autre. De plus, toute vectorisation ayant pour but de faire pénétrer dans la cellule non pas la cassette

d'expression, mais l'ARN actif est aussi envisagée. En particulier, tous les vecteurs et procédés facilitant l'introduction d'ARN dans les cellules seront utilisés.

5 C. PRODUCTION DE BANQUES D'EXPRESSION D'ARN ALEATOIRES

Une banque d'expression d'ARN aléatoires a été générée à partir de fragments d'ADN simple brin ayant une taille de 42 bases. Ces fragments se décomposent en trois parties : de 5' vers 3' on trouve une séquence connue de 8 bases (runA), puis une séquence aléatoire de 26 bases (synthétisée chimiquement de manière parfaitement aléatoire par un automate de synthèse d'ADN), pour finir, une deuxième séquence connue de 8 bases (runB). Les séquences runA et runB sont complémentaires. Dans un exemple spécifique, les fragments d'ADN simple-brin présentent la séquence 5' ATGAACGC (N)₂₆ GCGTTCAT 3', dans laquelle N représente toute base (A, T, G ou C). La partie aléatoire contient donc 26 positions variables.

Un oligonucléotide amorce complémentaire à runB a été utilisé comme amorce pour l'ARN polymérase de Kleenow, qui synthétise le brin d'ADN complémentaire à la totalité de la séquence. Dans le cas de l'exemple spécifique, l'amorce complémentaire à runB présente la séquence 5' ATGAACGC 3'.

A l'issue de cette étape de synthèse, on obtient une population d'ADN double brin présentant des extrémités franches appelée « banque de séquences aléatoires ».

25
$$\begin{array}{l} 5' - ATGAACGC (N)_{26} GCGTTCAT - 3' \\ 3' - TACTTGCG (N)_{26} CGCAAGTA - 5' \end{array} \quad \text{ADN double brin : « banque de séquences aléatoires »}$$

Cette banque de séquences aléatoires est ensuite clonée dans une cassette d'expression telle que décrite dans les exemples A et B ci-dessus. Selon le type d'expression recherché et/ou les besoins, les séquences aléatoires peuvent être insérées dans les deux types de cassettes dérivées de VA1 ou U6. Pour cela les plasmides ou vecteurs viraux (par exemple, pBabe

dans le cas des vecteurs rétroviraux) contenant les cassettes d'expression sont digérés par l'enzyme de restriction SrfI puis purifiés. La banque de séquences aléatoire (ADN double brin à bords francs) est clonée grâce à une ligation compétitive, en présence de l'enzyme de restriction SrfI. Statistiquement, chaque vecteur intègre donc un fragment d'ADN double
5 brin contenant une séquence aléatoire différente. La population hétérogène des plasmides ou vecteurs obtenus après insertion des séquences aléatoires est appelée «banque de plasmides aléatoires ou banque de vecteurs aléatoires». Ces banques constituent des banques d'expression d'ARN aléatoires, au sens de l'invention, utilisables directement pour la sélection in cellulo d'ARN actifs.

10

D. PRODUCTION D'UNE BANQUE DE RETROVIRUS ALEATOIRES

Dans un mode de mise en œuvre préféré, la banque d'expression utilisée est une banque
15 d'expression virale, notamment rétrovirale. Une telle banque peut être construite à partir d'une banque de vecteurs viraux aléatoires, comme décrit ci-après.

La banque de vecteurs aléatoires est introduite dans les cellules d'une lignée d'encapsulation (ici la lignée 293GP (Burns et al., 1993)) par transfection au phosphate de
20 calcium. La banque est simultanément transfectée avec un vecteur d'expression codant la glycoprotéine G d'enveloppe du virus de la stomatite vésiculeuse (VSV-G). Cette protéine est réputée être très efficace pour la fonctionnalité (stabilité, infectivité) des virus rétroviraux produits après transfection. Quelques jours après la transfection, les cellules produisent des rétrovirus pBabe recombinants. La «Banque de virus aléatoire» est alors
25 prélevée, purifiée et concentrée. Son titre infectieux ou MOI (nombre de particules rétrovirales recombinantes infectieuses par millilitre) peut alors être déterminé suivant les méthodes connues par l'homme de métier.

E. SELECTION IN CELLULO

- Les banques d'expression décrites dans les exemples C et D peuvent être utilisées pour la sélection in cellulo d'ARN actifs. Ainsi par exemple, avec la «Banque de virus aléatoires»,
- 5 il est possible de transférer les cassettes d'expression d'ARN aléatoires dans les cellules à étudier. L'objectif est d'obtenir une banque de cellules aléatoires dans laquelle chaque cellule exprime un, ou plusieurs, ARN aléatoire(s) différent(s). Les cellules utilisées peuvent être celles d'une lignée de référence, comme par exemple les cellules 293 (ATCC N° CRL 1573), les cellules Jurkat A3 (ATCC N° CRL-2570) ou les cellules HELA (ATCC
- 10 N° CCL 2). Dans une autre situation, les cellules à infecter ont une nature particulière dans le sens qu'elles présentent une activité spécifique sur laquelle on veut agir (croissance, différenciation, infection...). Dans tous les cas, on obtient une «banque de cellules aléatoires» qui peut être conservée et perpétuée à l'infini pour être utilisée dans différentes applications.
- 15 La «banque de virus aléatoires» permet de transduire les cellules d'une lignée quelconque ainsi que des cellules en culture primaire. Pour construire une «Banque de cellules aléatoires, standard ou spécifique», le principe est d'obtenir un maximum de cellules transduites ayant intégré chacune d'elle un nombre minimum de cassettes d'expression. L'objectif étant de se rapprocher de la valeur d'une seule séquence aléatoire intégrée par
- 20 génome cellulaire. L'infection des cellules se fait ainsi en utilisant un surnageant viral ayant une MOI inférieure à 1. La sélection des cellules transduites s'effectue grâce à la présence du gène de résistance à la puromycine dans les retrovirus pBabe recombinants. Après action de l'agent de sélection, la population de cellules obtenue est donc constituée par toutes les cellules ayant reçu une ou plusieurs copies du génome rétroviral et contenant donc une ou
- 25 plusieurs copies de cassette d'expression pol III codant chacune un ARN aléatoire différent.

Les cellules de la banque sont disponibles pour une première série de tests. La cible sur laquelle nous souhaitons agir grâce à l'ARN actif peut être connue (une protéine, un ARN, un ADN...) ou peut représenter une activité enzymatique, une voie métabolique, un

30 processus de prolifération ou de différenciation, une résistance à une drogue ou à un agent pathogène, etc.

Le test de sélection doit être adapté à chaque cas. Dans un mode préféré de l'invention, au sein de la banque de cellules aléatoires la sélection des cellules contenant un ARN agissant sur la cible est réalisée grâce à la sélection positive des cellules ayant acquis le phénotype requis et présentant un avantage sélectif. Dans un premier exemple, la sélection d'un ARN

5 actif contre un agent infectieux, tel qu'un virus cytopathogène (HIV), est réalisée grâce à la sélection positive des cellules devenues résistantes à la multiplication de ce virus. Dans un autre exemple, la sélection d'un ARN actif capable de protéger les cellules contre un processus de mort cellulaire programmée (apoptose) est réalisée au travers de la sélection des cellules devenues résistantes à l'ajout d'un signal déclenchant l'apoptose. La sélection

10 positive des cellules d'intérêts peut également comprendre une sélection directe des cellules par observation directe (surprolifération, changement de statut morphologique, différenciation altérée, expression d'un marqueur membranaire fluorescent...). Les colonies de cellules d'intérêts sont prélevées par microdissection à l'aide d'un automate permettant la microdissection au laser des groupes de cellules d'intérêts (Simone et al., 1998) ou bien

15 grâce à un tri positif basé sur la sélection des cellules d'intérêts marquées avec des anticorps. Après ce premier «tour de sélection», les cellules contenant potentiellement un ARN actif sont isolées : elles constituent la première génération de cellules sélectionnées. A ce stade, les cellules sélectionnées peuvent être amplifiées naturellement grâce à leur prolifération. Après une croissance suffisante, les cellules peuvent être clonées dans le cas

20 où elles apparaissent sous forme de clones indépendants comprenant chacun plusieurs milliers de cellules. Alternativement, les cellules peuvent être globalement réunies pour former une population de cellules du premier tour. Dans tous les cas, les cellules peuvent être conservées grâce à une congélation.

Cette première génération de cellules sélectionnées (population ou clones indépendants) est

25 directement utilisable pour effectuer un deuxième tour de sélection puis, plusieurs tours de sélection selon un mode itératif, avec la possibilité de faire varier les paramètres de sélection.

Dans certains cas, le mode de sélection et/ou le mode d'analyse des cellules sélectionnées peut nécessiter de travailler sur des cellules mortes car fixées par un agent fixateur comme

30 par exemple le formaldéhyde. C'est notamment le cas où le mode de sélection impose un marquage des cellules à l'aide d'anticorps. Dans ce cas, l'amplification naturelle des

cellules par croissance n'est pas possible et une étape supplémentaire d'amplification par PCR des cassettes contenant l'ARN actif est réalisée pour effectuer un deuxième tour de sélection.

L'ADN génomique des cellules issues du premier « tour de sélection » est extrait et purifié.

- 5 A partir de cet ADN, les techniques de biologie moléculaire permettent d'amplifier spécifiquement par PCR les séquences d'ADN correspondant aux cassettes d'expression d'ARN aléatoires contenues dans les cellules sélectionnées. On obtient ainsi une première génération de cassettes d'expression contenant des séquences potentiellement actives sur la cible. Cette première génération est dite « banque restreinte de cassettes du premier tour ».

10

F. SELECTION DES CASSETTES ACTIVES SUIVANT UN PROCEDE ITERATIF

- 15 La banque restreinte de cassettes du premier tour est traitée de la même manière que la banque aléatoire de départ. Elle est clonée dans le vecteur rétroviral pBabe au niveau du LTR 3'. Après cette nouvelle étape de clonage, on obtient une première banque restreinte de vecteurs viraux. Les formes infectieuses de ces vecteurs viraux sont obtenues à l'identique que précédemment par transfection des cellules d'encapsulation pour aboutir à une première banque restreinte de rétrovirus, utilisée pour infecter, à son tour, le type cellulaire étudié. La
- 20 MOI est à nouveau ajustée pour être inférieure à 1 copie de virus par cellule. Les cellules ainsi transduites sont à nouveau sélectionnées suivant les mêmes règles que celles qui régissent la sélection du premier tour, ou suivant d'autres paramètres. On aboutit ainsi à l'établissement d'une banque restreinte de cassettes du second tour. Ce deuxième tour de sélection permet d'enrichir la banque restreinte du premier tour en séquences actives.

- 25 La succession itérative des tours basés sur le principe d'une sélection pour le phénotype désiré permet d'enrichir, à chaque tour, la banque restreinte en séquence de cassettes actives. Lorsque le phénotype désiré semble être enfin stabilisé, c'est-à-dire lorsque que l'ensemble des cellules ayant reçu un vecteur viral recombinant présente le phénotype désiré (habituellement après 5 à 6 tours de sélection), il est considéré que la sélection des cassettes
- 30 dans les cellules est achevée.

G. IDENTIFICATION DES CASSETTES ACTIVES ET TEST DE LEUR FONCTIONALITE.

5 A partir des cellules sélectionnées, de préférence au cours du dernier tour de sélection, l'ADN génomique est extrait et purifié. Puis une amplification par PCR permet d'obtenir la dernière banque restreinte de cassettes actives. Le clonage de ces cassettes dans les vecteurs de clonage permet leur séparation physique grâce à la propagation de ces vecteurs dans des bactéries formant des colonies isolées. Le séquençage des cassettes retrouvées dans les
10 différentes colonies bactériennes (habituellement une trentaine de colonies sont analysées) permet de connaître les séquences aléatoires les plus fréquentes ou de déterminer un motif contenu dans les séquences aléatoires qui a été particulièrement conservé au cours du processus de sélection. La connaissance de ce motif permet de définir et de construire par biologie moléculaire une ou plusieurs cassettes contenant ce motif.

15 La ou les cassettes ainsi définies peuvent, à ce stade, être analysée(s) de manière individuelle afin de valider leur efficacité pour agir sur la cible. Pour cela, la cassette est clonée dans le vecteur pBabe, les vecteurs rétroviraux homogènes obtenus sont alors utilisés individuellement pour produire des virus recombinants qui servent à leur tour à infecter les
20 cellules étudiées. La comparaison de l'efficacité relative des différentes cassettes analysées permet de choisir la ou les cassettes les mieux adaptées pour agir sur la cible.

Ce procédé de criblage a été utilisé afin d'identifier des ARN actifs capables de rendre les cellules de la lignée Hela résistantes à l'apoptose cellulaire induite par la staurosporine (0,8
25 μ M – 6 heures). A l'issue d'un premier tour de sélection différents clones cellulaires résistants à la staurosporine ont été sélectionnés (Figure 12C) et les cassettes d'expression qu'ils contiennent ont été identifiées : clones 2, 5, 9, 11, 13, 14, 15, 16, C, J, L et N. La validation des ARN actifs ainsi identifiés a été effectuée selon le mode indiqué ci-dessus dans les cellules Hela ainsi que dans les cellules Jurkat (avec une multiplicité d'infection
30 inférieure à 1). Dans chacune des populations cellulaires, le niveau d'expression des ARN actifs a été évalué par Northern blot (Figure 12D). Dans cette figure, nous observons que le

niveau d'expression de chacun des ARN est comparable d'une lignée cellulaire à l'autre (Hela versus Jurkat). De nouveaux tests de résistance à l'apoptose induite par la staurosporine ont été effectués sur les populations de cellules Hela ou Jurkat exprimant chacune d'elle l'un des ARN actifs. Sur les 12 ARN étudiés, seuls certains présentent une
5 activité significative de résistance à la Staurosporine, les autres sont de faux positifs. Les résultats obtenus dans ces deux lignées cellulaires indiquent que le taux de résistance des cellules à l'apoptose varie en fonction de l'ARN exprimé. Parmi les ARN sélectionnés, les clones 5, 9 13, 15 et 16 présentent une activité significative dans les cellules Jurkat (l'ARN clone 9 étant le plus actif) alors que seul le clone 9 présente une activité anti-apoptotique
10 dans les cellules Hela (Figure 12E). Le recours à un système inductible tel que le système VAI peut faciliter les différentes étapes décrites dans ce procédé afin de valider les ARN actifs directement dans les cellules dans lesquelles ils ont été sélectionnés (élimination directe des faux positifs, Figure 15 Panneau A)). Le système VAI est dans cet objectif validé dans les cellules de la lignée Hela T-Rex (invitrogen ref: R714-07) qui expriment
15 constitutivement le transgène TetR du répresseur bactérien du gène de la Tétracycline (Figure 7C).

H - CONSTRUCTION DE LIBRAIRIES DE CASSETTES D'EXPRESSION 20 ADAPTÉES À L'EXPRESSION IN VITRO

Une librairie de cassettes aléatoires a été générée in vitro à partir de fragment d'ADN simple brin en trois étapes (Figure 13A).

La première étape consiste à hybrider et allonger deux brins d'ADN afin de reconstituer
25 une librairie de fragments du gène VA dans lequel est insérée une séquence aléatoire.

Ce fragment sert alors de matrice à une réaction de polymérisation en chaîne qui permet d'obtenir la librairie de cassette d'expression VA aléatoire dans sa totalité (2^{ème} étape).

La 3^{ème} étape consiste à ajouter, en amont des cassettes d'expression VA aléatoire, les séquences d'un promoteur utilisable pour effectuer une transcription in vitro.

A partir de cette librairie de transcription VA aléatoire, une étape de transcription in vitro permet l'obtention de la librairie d'ARN VA aléatoire.

Dans un exemple spécifique, les fragments d'ADN simple brin de la première étape sont les suivants :

brin sens : banque Sens

5'GCGACCGGGGTTCTGAACCCCGGAATAACTCTATCAATGATATGCCCAGCCC3'

brin antisens : banque Antisens

5'-CGAACTTCTTGATGCCCTGCCC(N)₃₀3'

10 dans lequel N représente une base aléatoire : A ou T ou G ou C

Dans cet exemple, la séquence aléatoire est représentée par un fragment de 30 bases flanqué en amont et en aval de deux séquences constantes de 5 bases, capables de former une structure en double hélice dans la molécule d'ARN finale. Cette structure se trouve positionnée à un niveau qui équivaut à l'insertion des séquences aléatoires dans le site SrfI de la cassette d'expression inductible VAIo (voir figure 7). Après une étape d'hybridation, les deux oligonucléotides sont allongés par l'ADN polymérase fragment de Kleenow qui produit le fragment VAIo aléatoire double brin (Figure 13 A). La deuxième étape consiste à construire à partir du fragment VAIo aléatoire double brin la librairie de cassettes d'expression inductible VAIo aléatoire de pleine taille. Cette étape a été réalisée par une réaction de polymérisation en chaîne à l'aide des oligonucléotides suivants :

oligo sens : VAI5'

5'GGGCACTCTTCCGTGGTCTGGTGGATAAACTCTATCATTGATAGAGTTATGCG
ACCGGGGTTCTGAACCCCGG 3'

oligo antisens : VAI end banque

25 5-AAAAGGAGCGCTCCCCGTTGTCTGACGTCGAACTTCTTGATGCCCTGCCC-3'

utilisant une Taq DNA polymérase (Figure 13 A). Au cours de la troisième étape, le promoteur du bactériophage T7 a été rajouté en amont de chaque cassette d'expression inductible VAIo aléatoire pour générer la librairie de cassettes de transcription inductibles VAIo aléatoires. Cette étape a été réalisée par réaction de polymérisation en chaîne en

30 utilisant les oligonucléotides

VApT7 :

5' -AAATTAATACGACTCACTATAGGGCACTCTTCCGTGGTCTGG -3' en amont

et VA end :

5' -AAAAGGAGCGCTCCCCGTTG -3' en aval.

- 5 La librairie de cassettes de transcription T7 VAIo aléatoire ainsi obtenue a été utilisée pour générer in vitro une librairie d'ARN VAIo aléatoire grâce à l'utilisation de l'ARN polymérase du bactériophage T7 (Figure 13 B). Dans ces expériences plusieurs dizaines de microgrammes d'ARN ont été obtenues.

10

I – CRIBLAGE IN VITRO D'ARN ACTIFS ET OBTENTION DE LIBRAIRIES RESTREINTES DE CASSETTES D'EXPRESSION

- La librairie d'ARN VA aléatoires est utilisée pour sélectionner des ARN capables de lier in vitro un substrat spécifique. Les ARN actifs sélectionnés par toute méthode adaptée, connue de l'homme du métier, servent de matrice pour générer par une étape de transcription réverse suivie d'une étape d'amplification (réaction de RT-PCR) les cassettes d'expression correspondantes. Ce mélange de cassettes d'expression d'ARN actifs est alors cloné dans un vecteur adapté au transfert de gènes puis utilisé comme matériel de départ pour la sélection in cellulo d'ARN actifs (librairie de cassettes d'expression d'ARN aléatoires enrichie en ARN actifs).
- 15 20

- Dans un exemple spécifique, les ARN VAIIV Srf, VA TAR* ou VAIo aléatoires ont été utilisés pour obtenir les cassettes de transcription correspondantes. La réaction de RT-PCR a été effectuée en présence des amorces adaptées : VApT7 et VAend (Figure 13 C). D'autre part, afin de montrer que le produit de la réaction de RT-PCR (est bien représentatif de la diversité du substrat, des mélanges d'ARN ont été utilisés comme matrice. La figure 13 C montre que dans les conditions dans lesquelles trois ARN matrice sont mélangés (ARN VAIIV Srf ARN VATAR* et ARN VAIo aléatoires) de façon équimoléculaire (1/3 ; 1/3 ; 1/3) la réaction de RT-PCR produit trois cassettes dont les quantités respectives reflètent les quantités initiales de chacun des ARN substrat. Ainsi dans les conditions où la
- 25 30

librairie d'ARN VAiO aléatoire est utilisée comme substrat, les produits de la réaction de RT-PCR sont représentatifs d'une librairie de cassettes d'expression.

J – EXPRESSION D'ARN ACTIFS DE SEQUENCE DETERMINEE

5

Les constructions de l'invention peuvent également être utilisées pour tester des séquences actives définies, et/ou exprimer de telles séquences dans des tissus biologiques.

10 Dans ce cas, la séquence active à insérer se présente sous la forme d'ADN double brin dont la séquence a été choisie pour générer un ARN actif efficace (de type antisens, ribozyme, ARN interférant (siARN, miARN ou leurs précurseurs) ou ARN aptamère). L'ADN double brin peut être obtenu par différentes techniques comme l'hybridation entre deux oligonucléotides complémentaires, la purification de fragments de restriction, la copie d'une matrice par PCR, etc.

15 Un vecteur tel que décrit dans l'exemple B, contenant une cassette d'expression, est digéré par l'enzyme de restriction adaptée et le clonage de la séquence active s'effectue dans ce site de restriction par des techniques classiques de clonage.

Des vecteurs peuvent ainsi être produits, qu'il s'agisse de vecteurs plasmidiques ou viraux, par exemple. D'autre part, des virus recombinants peuvent également être générés, comme
20 décrit dans l'exemple C. Les rétrovirus recombinants produits servent alors à infecter les cellules dont on souhaite altérer le phénotype. L'infection est faite préférentiellement avec une forte multiplicité d'infection afin d'intégrer un nombre élevé de cassettes d'expression d'ARN actif dans le génome cellulaire. En effet, l'activité de la séquence active est fortement dépendante de son niveau d'expression (figure 11). La présence du gène de
25 résistance à la puromycine dans les rétrovirus pBabe permet une sélection rapide des cellules qui ont été infectées et contenant la séquence transduite.

Ces vecteurs peuvent également être purifiés et conditionnés dans tout véhicule ou excipient acceptable, pour produire des compositions administrables, par exemple chez des
30 organismes mammifères, notamment humain.

REFERENCES

- Barcellini, C. S., Fenard, D., Bertrand, E., Singer, R. H., Lefebvre, J. C., and Doglio, A. (1998). 3'-End modification of the adenoviral VA1 gene affects its expression in human cells: consequences for the design of chimeric VA1 RNA ribozymes [In Process Citation]. *Antisense Nucleic Acid Drug Dev* 8, 379-90
- 5 379-90.
- Bertrand, E., Castanotto, D., Zhou, C., Carbonnelle, C., Lee, N. S., Good, P., Chatterjee, S., Grange, T., Pictet, R., Kohn, D., Engelke, D., and Rossi, J. J. (1997). The expression cassette determines the functional activity of ribozymes in mammalian cells by controlling their intracellular localization. *Rna* 3, 75-88
- 10 75-88.
- Bertrand, E., Gwizdek, C., Fenard, D., and Doglio, A. (1999). L'ARN molécule thérapeutique ? vers une conception rationnelle du développement des ARN antisens, des ribozymes et des aptamères d'ARN. *Médecine/Sciences* 15, 677-681.
- 15
- Burns, J. C., Friedmann, T., Driever, W., Burrascano, M., and Yee, J. K. (1993). Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci U S A* 90, 8033-7.
- 20
- Cagnon, L., Cucchiaroni, M., Lefebvre, J. C., and Doglio, A. (1995). Protection of a T-cell line from human immunodeficiency virus replication by the stable expression of a short antisense RNA sequence carried by a shuttle RNA molecule. *J Acquir Immune Defic Syndr Hum Retrovirol* 9, 349-58
- 25 349-58.
- Charneau, P., Mirambeau, G., Roux, P., Paulous, S., Buc, H., and Clavel, F. (1994). HIV-1 reverse transcription. A termination step at the center of the genome. *J Mol Biol* 241, 651-62.
- 30
- Famulok, M., and Verma, S. (2002). In vivo-applied functional RNAs as tools in proteomics and genomics research. *Trends Biotechnol* 20, 462-6.
- 35
- Gmeiner, W. H. (2002). The structure and dynamics of the U4/U6 snRNP: implications for pre-mRNA splicing and use as a model system to investigate the RNA-mediated effects of (5F)Ura. *J Biomol Struct Dyn* 19, 853-62.
- 40
- Gold, L. (1995). The SELEX process: a surprising source of therapeutic and diagnostic compounds. *Harvey Lect* 91, 47-57
- 47-57.

- Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA* 89, 5547-5551.
- 5 Gwizdek, C., Bertrand, E., Dargemont, C., Lefebvre, J. C., Blanchard, J. M., Singer, R. H., and Doglio, A. (2001). Terminal minihelix, a novel RNA motif that directs polymerase III transcripts to the cell cytoplasm. Terminal minihelix and RNA export. *J Biol Chem* 276, 25910-8.
- 10 Hermann, T., and Patel, D. J. (2000). Adaptive recognition by nucleic acid aptamers. *Science* 287, 820-5.
- Hillen, W., and Berens, C. (1994). Mechanisms underlying expression of Tn10 encoded tetracycline resistance. *Annu Rev Microbiol* 48, 345-69.
- 15 Mathews, M. B., and Shenk, T. (1991). Adenovirus-associated RNA and translation control. *J. Virol.* 65, 5657 - 5662.
- McManus, M. T., and Sharp, P. A. (2002). Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet* 3, 737-47.
- 20 Medina, M. F., and Joshi, S. (1999). RNA-polymerase III-driven expression cassettes in human gene therapy. *Curr Opin Mol Ther* 1, 580-94.
- Morgenstern, J. P., and Land, H. (1990). Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line. *Nucleic Acids Res* 18, 3587-96.
- 25 Paule, M. R., and White, R. J. (2000). Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res* 28, 1283-98.
- 30 Scherr, M., Morgan, M. A., and Eder, M. (2003). Gene Silencing Mediated by Small Interfering RNAs in Mammalian Cells. *Curr Med Chem* 10, 245-56.
- 35 Simone, N. L., Bonner, R. F., Gillespie, J. W., Emmert-Buck, M. R., and Liotta, L. A. (1998). Laser-capture microdissection: opening the microscopic frontier to molecular analysis. *Trends Genet* 14, 272-6.
- 40 Yamamoto, R., Katahira, M., Nishikawa, S., Baba, T., Taira, K., and Kumar, P. K. (2000). A novel RNA motif that binds efficiently and specifically to the Ttat protein of HIV and inhibits the trans-activation by Tat of transcription in vitro and in vivo. *Genes Cells* 5, 371-88.

REVENDICATIONS

1. Procédé de sélection in cellulo d'ARN actifs capables de conférer à une cellule un phénotype désiré, comprenant :
 - 5 a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression dérivée d'un gène VA d'un adénovirus exprimant un ARN structural aléatoire distinct,
 - b) la mise en contact de ladite banque ou d'une partie de cellule-ci avec une population
10 de cellules dans des conditions permettant l'infection desdites cellules par lesdits rétrovirus recombinants,
 - c) la sélection des cellules possédant le phénotype désiré, et
 - d) l'identification de la ou des cassette(s) contenue(s) dans la ou lesdites cellules, ou de l'ARN actif qu'elles expriment.
- 15 2. Procédé selon la revendication 1, caractérisé en ce que la cassette comprend la séquence d'un gène VA1 d'adénovirus délété de tout ou une partie fonctionnelle de la boucle IV, dans laquelle la séquence de l'ARN structural aléatoire est insérée.
- 20 3. Procédé selon la revendication 1 ou 2, caractérisé en ce que la double-hélice terminale du gène VA est altérée.
4. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que le gène VA comprend une séquence conférant un caractère inductible, de préférence entre les
25 boîtes A et B et/ou en amont du gène, plus préférentiellement en remplacement de tout ou partie des séquences natives.
5. Procédé selon la revendication 4, caractérisé en ce que la séquence conférant un caractère inductible est une séquence conférant une sensibilité à un facteur agissant en trans, de
30 préférence un site de liaison d'un facteur transcriptionnel ou d'un répresseur.

6. Procédé selon la revendication 5, caractérisé en ce que la séquence conférant un caractère inductible est représentée par une ou plusieurs séquences opératrices d'un promoteur bactérien régulé, de préférence du promoteur de la tétracycline.

5 7. Procédé selon la revendication 4, caractérisé en ce que la cassette d'expression inductible est choisie parmi les cassettes OVAi (SEQ ID NO: 16), VAiO (SEQ ID NO: 15) et OVAiO (SEQ ID NO: 17).

10 8. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que la cassette comprend en outre un second promoteur transcriptionnel, en amont du promoteur VA.

9. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que l'ARN actif est un ARN structural.

15

10. Procédé selon l'une quelconque des revendications 1 à 8, caractérisé en ce que l'ARN actif est un ARN interférant (siARN, miARN ou leurs précurseurs) ou antisens.

20 11. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que la banque est mise en contact avec la population de cellules dans des conditions permettant le transfert d'un nombre restreint de cassettes par cellule, de préférence à une MOI inférieure à 1.

25 12. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que la banque d'acides nucléiques est préparée par :

- synthèse d'une banque d'ADN simple-brins comprenant une région aléatoire encadrée par une ou deux régions de séquence définie,
 - synthèse d'un deuxième brin au moyen d'une ADN polymérase et en présence d'une amorce complémentaire de la séquence définie du premier brin, ou d'une partie de
- 30 celle-ci, pour produire une banque d'ADN double-brins comprenant une région aléatoire,

- clonage de la banque d'ADN double-brins dans un vecteur rétroviral, sous contrôle d'un promoteur dérivé d'un gène VA d'un adénovirus, et
- transfection dudit vecteur dans une lignée cellulaire d'encapsidation, pour produire une banque de rétrovirus recombinants.

5

13. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que la banque mise en œuvre dans l'étape a) est une banque aléatoire restreinte codant des ARN aléatoires pré-sélectionnés pour une ou des propriétés particulières, de préférence pour leur capacité de liaison et/ou pour leur affinité pour une cible, pour la présence d'un motif structural et/ou d'une séquence spécifique.

10

14. Procédé selon l'une quelconque des revendications 1 à 12, caractérisé en ce que la banque mise en œuvre dans l'étape a) est une banque de mutants aléatoires d'un gène cible particulier.

15

15. Procédé selon la revendication 13, comprenant :

1a) la préparation d'une banque de rétrovirus recombinants comprenant une cassette d'expression comprenant une séquence nucléique codant un ARN aléatoire placée sous le contrôle d'un promoteur dérivé d'un gène VA d'un adénovirus, les ARN aléatoires codés ayant été pré-sélectionnés in vitro pour leur capacité à lier et/ou pour leur affinité pour une cible d'intérêt,

20

1b) la mise en contact de cette banque ou d'une partie de cette banque, avec une population de cellules dans des conditions permettant le transfert d'acide nucléique dans lesdites cellules,

25

2) la sélection d'une ou plusieurs cellules possédant le phénotype désiré, et

3) l'identification de la ou des cassette(s) contenue(s) dans la ou lesdites cellules, ou de(s) l'ARN(s) actif(s) qu'elles expriment.

16. Procédé selon la revendication 15, caractérisé en ce que la banque de rétrovirus recombinants codant des ARN aléatoires présélectionnés est préparée par:

30

- synthèse in vitro d'une banque de cassettes d'expression (ADN db) comprenant chacune une séquence nucléique codant un ARN aléatoire placée sous le contrôle d'un promoteur dérivé d'un gène VA d'un adénovirus et, le cas échéant, d'un promoteur supplémentaire;
- expression in vitro de la banque de cassettes ou d'une partie de celle-ci, produisant in vitro
- 5 une banque d'ARN aléatoires;
- sélection in vitro des ARN pour leur capacité de liaison ou pour leur affinité pour une cible donnée; et
- production d'une banque restreinte de rétrovirus recombinants comprenant des cassettes d'expression d'ARN ainsi sélectionnés.

10

17. Procédé selon la revendication 16, caractérisé en ce que l'expression in vitro de la banque de cassettes est réalisée en deux étapes, une étape de production de cassettes de transcription et une étape de transcription proprement dite.

- 15 18. Procédé de sélection d'ARN actifs, comprenant (i) la synthèse in vitro d'une banque de cassettes d'expression (ADN db) codant pour des ARN aléatoires sous contrôle d'un promoteur transcrit par l'ARN polymérase III dérivé d'un gène VA d'un adénovirus, (ii) l'expression in vitro de la banque de cassettes ou d'une partie de celle-ci, produisant in vitro une banque d'ARN aléatoires, et (iii) la sélection in vitro des ARN pour leur capacité de
- 20 liaison ou pour leur affinité pour une cible donnée.

25

19. Procédé selon la revendication 18, comprenant en outre une étape (iv) de production d'une banque de restreinte de rétrovirus recombinants comprenant des cassettes d'expression d'ARN sélectionnées à l'étape (iii).

20. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que l'ADN des cassettes d'expression des cellules sélectionnées est amplifié pour produire une banque restreinte, et en ce que les étapes b) et c) du procédé sont répétées une fois au moins avec ladite banque restreinte.

30

21. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que la population de cellules comprend des cellules animales (par exemple de mammifère), de levure, de plante, d'insecte ou bactériennes.

5 22. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que le phénotype désiré est choisi parmi une capacité ou une incapacité de croissance, d'apoptose, de différenciation, de migration, de résistance à un agent toxique, de résistance à un agent infectieux, ou d'action métabolique.

10 23. Procédé selon l'une des revendication 1 à 21, caractérisé en ce que le phénotype désiré est l'expression d'une cible biologique.

24. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que la population de cellules comprend des cellules infectées par un virus et le phénotype désiré
15 est la résistance audit virus.

25. Procédé selon l'une quelconque des revendications 1 à 23, caractérisé en ce que la population de cellules comprend des cellules tumorales et le phénotype désiré est la perte de tumorigénicité.

20

26. Procédé selon l'une quelconque des revendications 1 à 23, caractérisé en ce que la population de cellules comprend des cellules bactériennes et le phénotype désiré est la sensibilité à un agent toxique.

25 27. Procédé selon l'une quelconque des revendications 4 à 26, caractérisé en ce que l'activité de l'ARN est sélectionnée en cultivant les cellules en condition d'induction de l'expression et, de préférence, validée en comparant l'état induit et l'état réprimé.

28. Procédé de sélection d'ARN actifs capables de conférer à une cellule un phénotype
30 désiré, comprenant :

- a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette d'expression dérivée de la séquence d'un gène VA d'un adénovirus exprimant un ARN aléatoire distinct,
- 5 b) la mise en contact de ladite banque, ou d'une partie de celle-ci, avec une population de cellules dans des conditions permettant l'infection desdites cellules par lesdits rétrovirus recombinants,
- c) la sélection de cellules possédant le phénotype désiré,
- d) l'extraction ou l'amplification de la séquence de cassettes contenues dans lesdites
- 10 cellules,
- e) le clonage des séquences obtenues en d) dans un vecteur pour générer une banque restreinte et,
- f) la répétition une fois au moins des étapes b) à d) avec ladite banque restreinte.
- 15 29. Procédé selon la revendication 28, dans lequel ladite banque d'acides nucléiques de l'étape a) code des ARN aléatoires pré-sélectionnés pour une propriété particulière, de préférence pour leur capacité à lier, in vitro, une cible d'intérêt.
30. Procédé de sélection d'ARN actifs sur une cible biologique déterminée, comprenant :
- 20 a) la fourniture d'une banque d'acides nucléiques comprenant une pluralité de vecteurs comprenant des cassettes d'expression distinctes placées sous le contrôle d'un promoteur transcrit par l'ARN polymérase III et comprenant chacune une séquence nucléique codant un ARN contraint ou prédéfini pour agir sur une cible donnée,
- b) la mise en contact de ladite banque, ou d'une partie de celle-ci, avec une population
- 25 de cellules exprimant ou contenant la cible biologique, dans des conditions permettant le transfert d'acide nucléique dans lesdites cellules,
- c) la sélection de cellules possédant le phénotype désiré,
- d) l'extraction ou l'amplification de la séquence de cassettes contenues dans lesdites cellules et, de manière facultative,
- 30 e) le clonage des séquences obtenues en d) dans un vecteur pour générer une banque restreinte et,

f) la répétition une fois au moins des étapes b) à d) avec ladite banque restreinte.

31. Banque d'acides nucléiques, caractérisée en ce qu'elle comprend une pluralité d'espèces de rétrovirus recombinants, chaque espèce de rétrovirus comprenant une cassette
5 d'expression dérivée d'un gène VA d'un adénovirus exprimant un ARN structural aléatoire distinct.

32. Cassette d'expression d'un ARN, caractérisée en ce qu'elle comprend une séquence codant ledit ARN insérée dans un promoteur dérivé d'un gène VA d'un adénovirus, ledit
10 promoteur comprenant en outre une séquence conférant un caractère inductible.

33. Cassette selon la revendication 32 ou banque selon la revendication 31, caractérisée en ce que la double-hélice terminale du gène VA est altérée.

15 34. Cassette selon la revendication 32 ou 33, ou banque selon la revendication 31, caractérisée en ce que l'ARN est un ARN structural aléatoire ou de séquence définie.

35. Cassette selon l'une des revendications 32-34, ou banque selon la revendication 31, caractérisée en ce que la cassette comprend en outre un second promoteur transcriptionnel,
20 localisé en amont du promoteur VA.

36. Vecteur comprenant une cassette selon l'une des revendications 32 à 35.

37. Cellule comprenant une cassette selon l'une des revendications 32 à 35 ou un vecteur
25 selon la revendication 36.

38. Composition pharmaceutique ou formulation, caractérisée en ce qu'elle comprend une cassette selon l'une des revendications 32 à 35, un vecteur selon la revendication 36, une
cellule selon la revendication 37 ou une cassette d'expression d'un ARN actif identifié ou
30 produit par un procédé selon l'une quelconque des revendications 1 à 30, ou un ARN actif constitué uniquement par le motif actif isolé au sein de ladite cassette.

1	GGGCACTCTT	CCGTGGTCTG	GTGGATAAAT	TCGCAAGGGT
41	ATCATGGCGG	ACGACCGGGG	<u>TTCGAACCCC</u>	GGATCCGGCC
81	GTCCGCCGTG	<u>ATCCATGCGG</u>	<u>TTACCGCCCG</u>	<u>CGTGTCGAAC</u>
121	CCAGGTGTGC	GACGTCAGAC	AACGGGGGAG	CGCTCCTTTT

FIGURE 1A

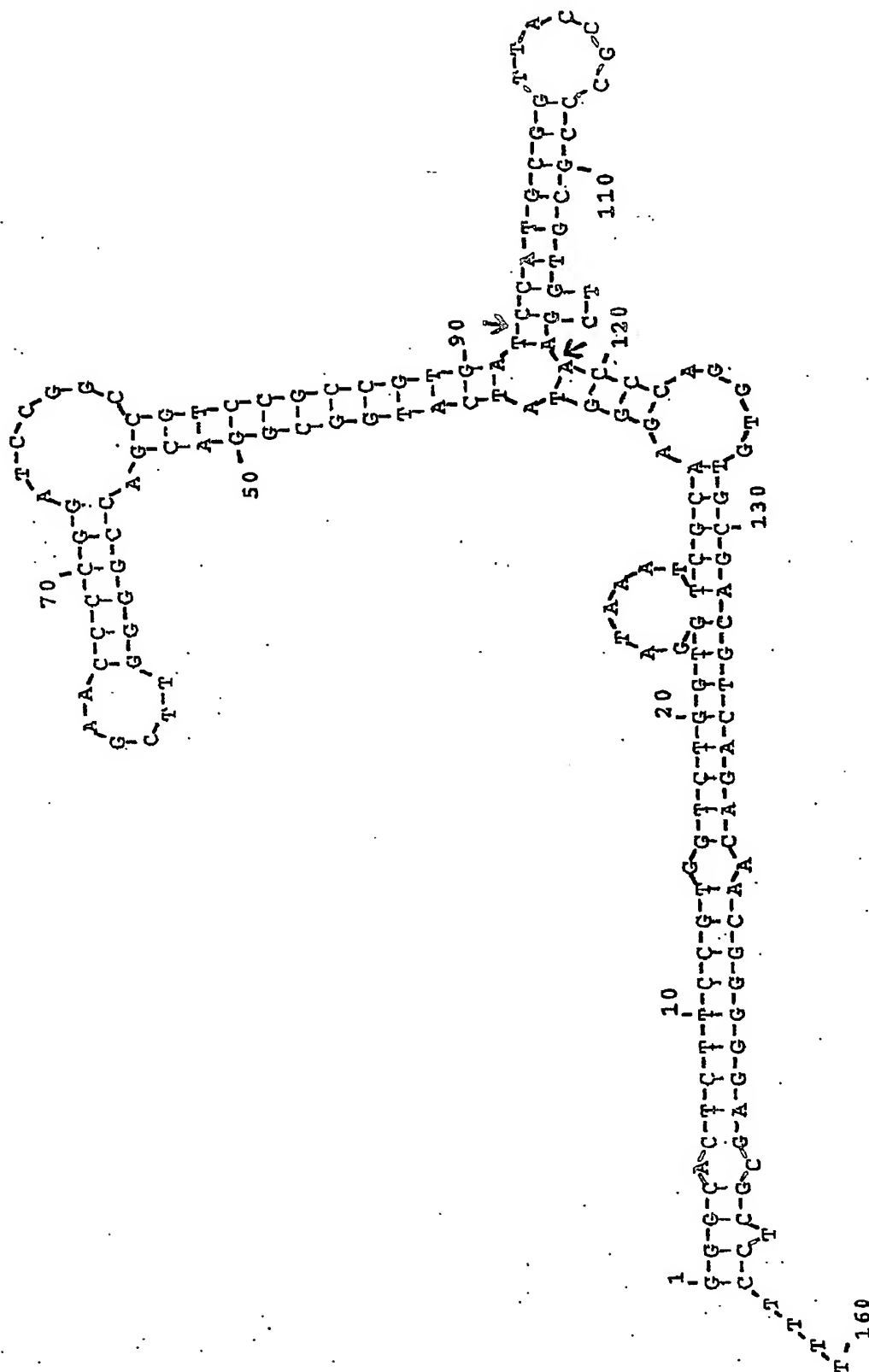


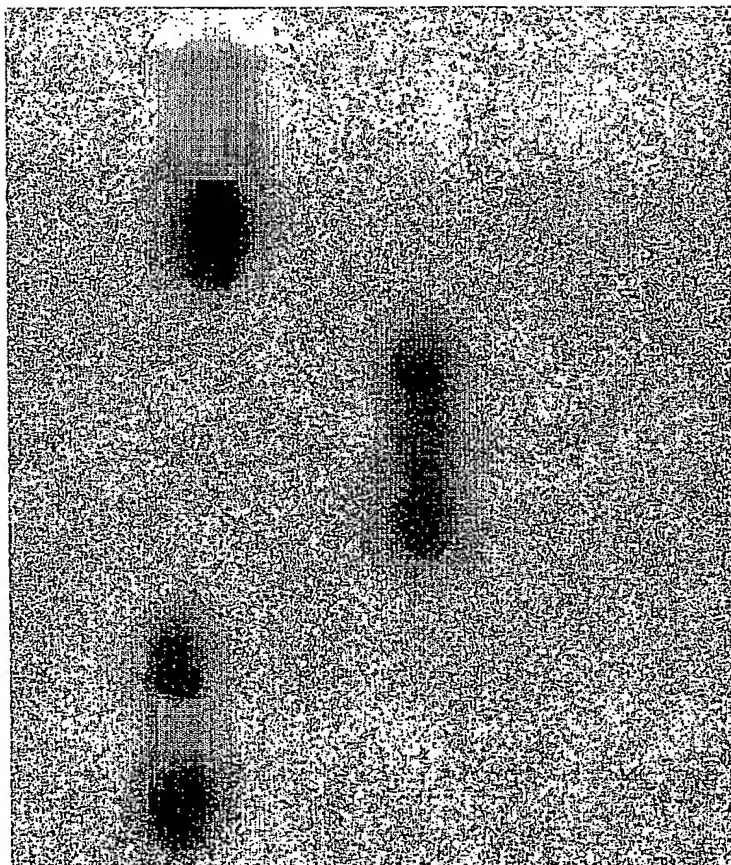
FIGURE 1B : Structure secondaire de l'ARN VAI

3/29

5'-ACGAAGCTTGATCCCGTATGCCGGTCGATCGCTTCGA-3'

FIGURE 2A

4/29

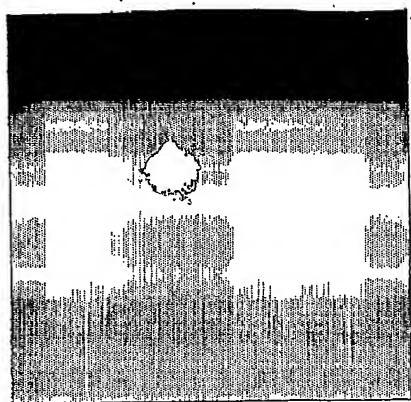


VATAR* VAΔIVSrf VA1

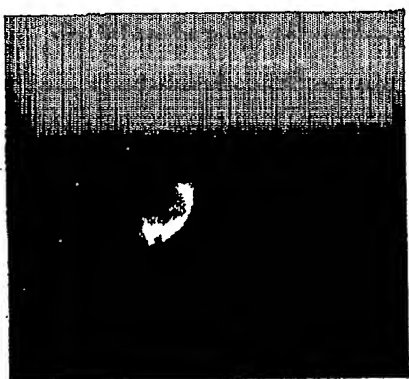
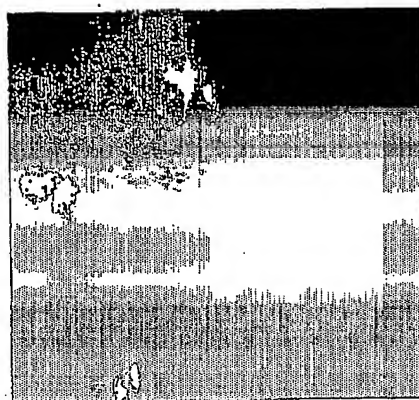
Figure 2B

BEST AVAILABLE COPY

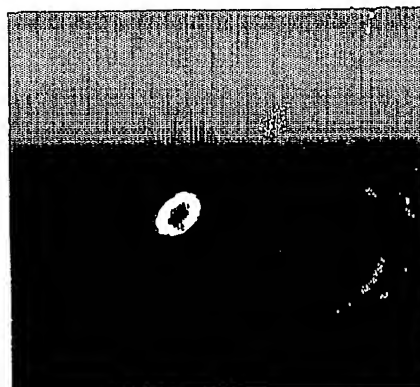
5/29



VA 1



VAA IV
Srf



VA
TAR*



Sonde spécifique

Coloration DAPI

Figure 2C : Localisation cellulaire des ARN
VA1, VAA IV Srf et VATAR*

BEST AVAILABLE COPY

6/29

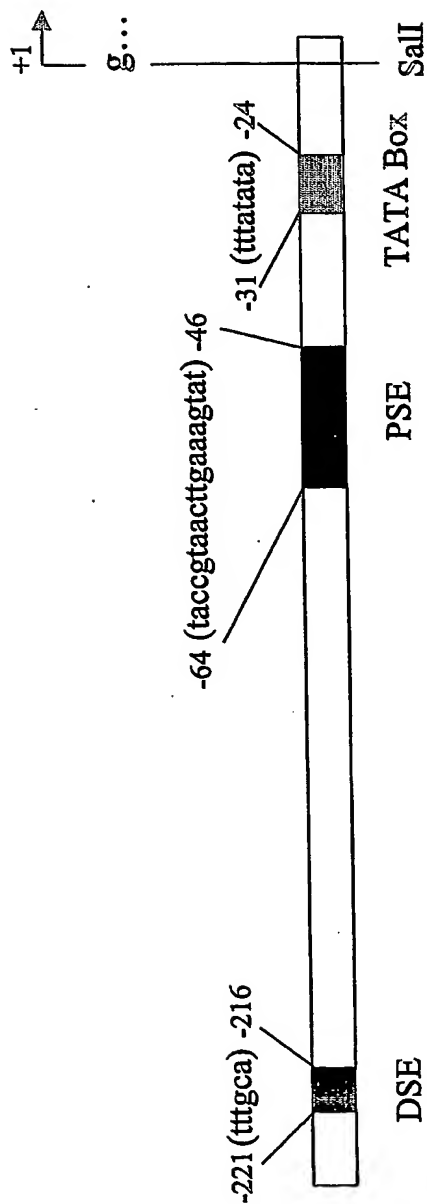


FIGURE 3

BEST AVAILABLE COPY

7/29

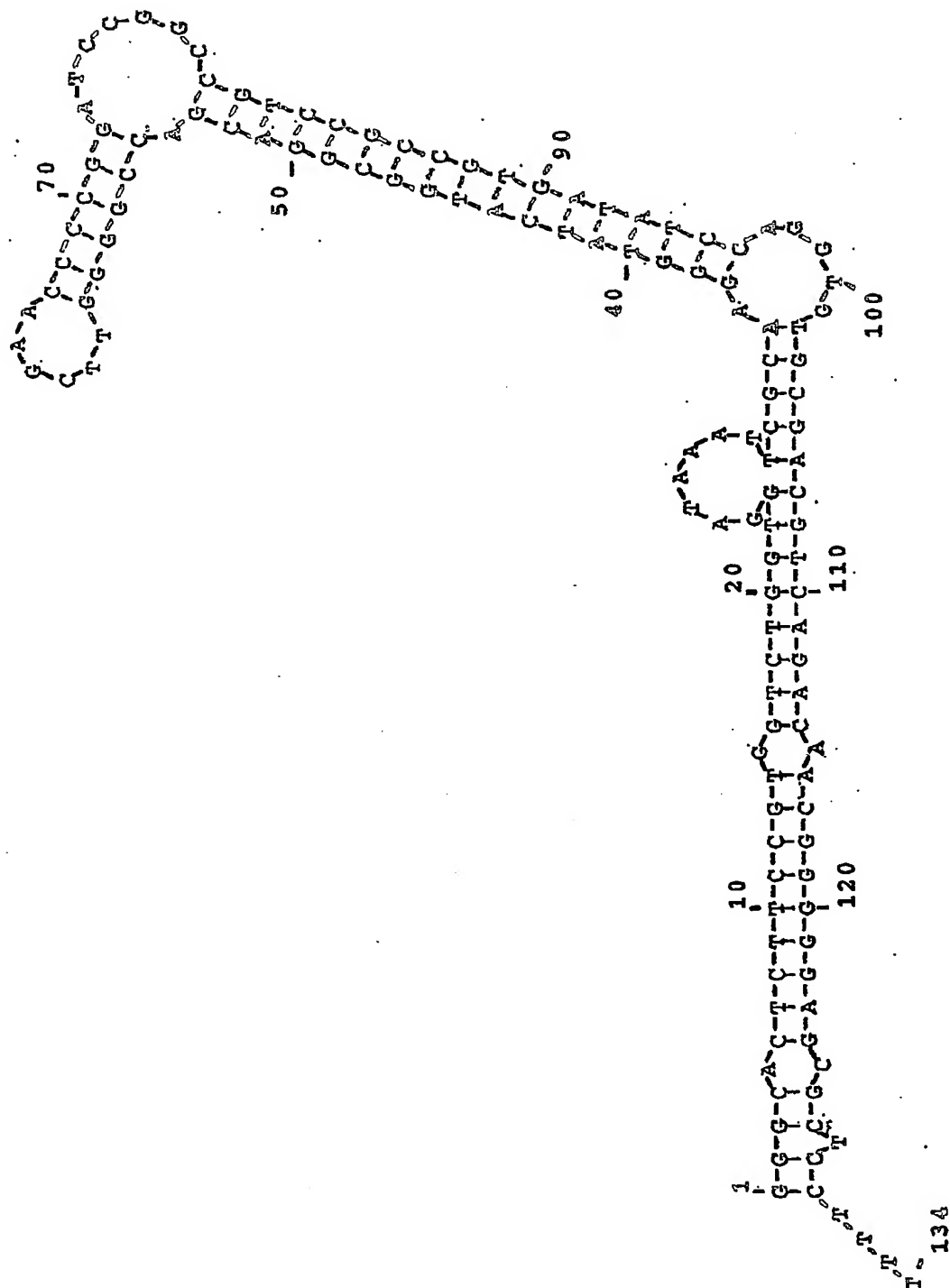


FIGURE 4 : Structure secondaire de l'ARN VAAIV

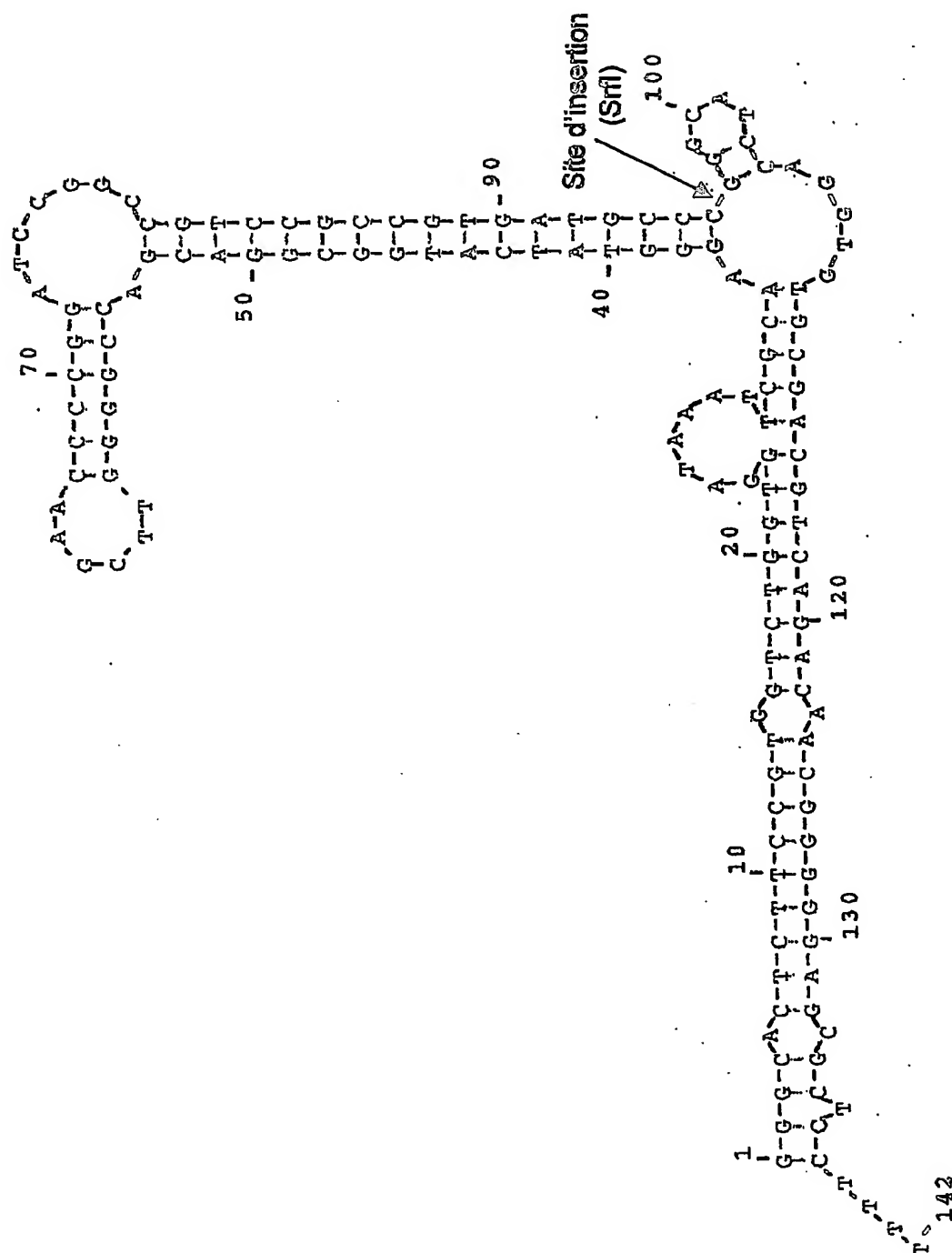


FIGURE 5 : Structure secondaire de l'ARN VAAIVS^{sf}

9/29

```
1  GGGCACTCTT  CCGTGGTCTG  GTGGATAAAT  TCGCAAGGGT
41 ATCATGGCGG  ACGACCGGGG  TTCGAACCCC  GGATCCGGCC
81 GTCCGCCGTG  ATGCCCGGGC  ATCCAGGTGT  GCGACGTCA a
121 t a AACGGGGG  AGCGCCCTTT  T
```

FIGURE 6A

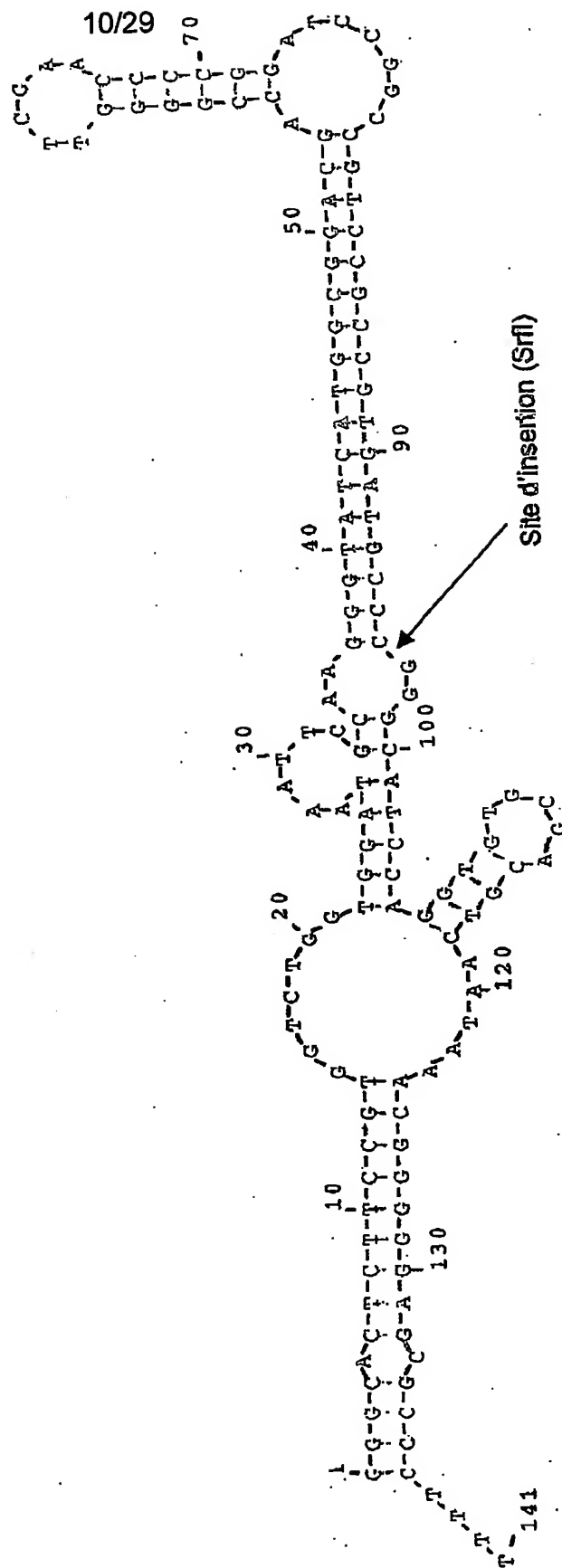
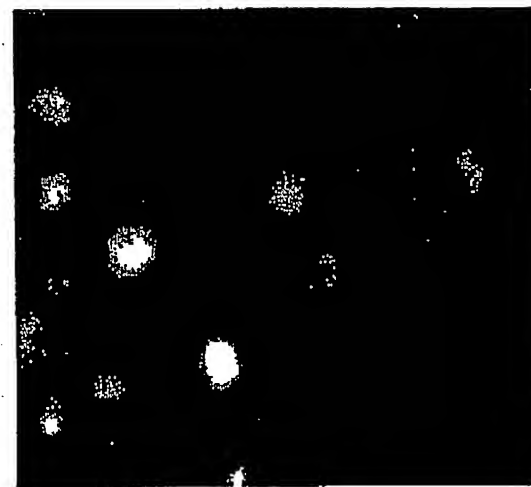
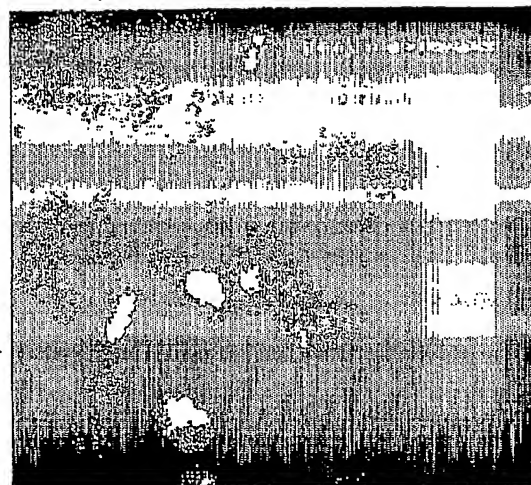
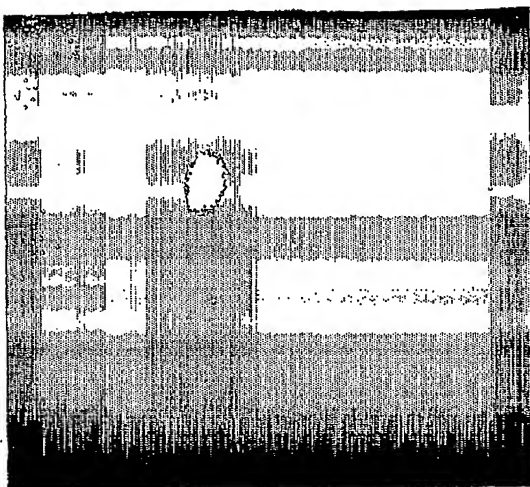


FIGURE 6B : Structure secondaire de l'ARN nVAΔIVS⁸

11/29



Sonde spécifique

Coloration DAPI

Figure 6C : Localisation cellulaire des ARN $nVAA\Delta IV$ Srf

12/29

Gène VAiO

```

1  GGGCACTCTT  CCGTGGTCTG  GTGGATAAAC  TCTATCATTG
41  ATAGAGTTAT  GCGACCGGGG  TTCGAACCCC  GGAATAACTC
81  TATCAATGAT  ATGCCCAGGC  ATCAAGAAGT  TCGACGTCAG
121 ACAACGGGGG  AGCGCTCCTT  TT

```

Gène OVAi

```

-50
-40 TGATAGAGTT  ATAGACCGTG  CAAAAGGAGA  ACTCTATCAT
1  GGGCACTCTT  CCGTGGTCTG  GTGGATAAAC  GCCTGTAAGC
41  TAGAGTTATG  CGACCGGGGT  TCGAACCCCG  TCTATCATTG
81  ATCAATGATA  TGCCCAGGCA  TCAAGAAGTT  GAATAACTCT
121 CAACGGGGGA  GCGCTCCTTT  T  CGACGTCAGA

```

Gène OVAiO

```

-50
-40 TGATAGAGTT  ATAGACCGTG  CAAAAGGAGA  ACTCTATCAT
1  GGGCACTCTT  CCGTGGTCTG  GTGGATAAAC  GCCTGTAAGC
41  ATAGAGTTAT  GCGACCGGGG  TTCGAACCCC  TCTATCATTG
81  TATCAATGAT  ATGCCCAGGC  ATCAAGAAGT  GGAATAACTC
121 ACAACGGGGG  AGCGCTCCTT  TT  TCGACGTCAG

```

FIGURE 7A

13/29

VAi up (90 bases)

5'- CTGTAAGCGG GCACTCTTCC GTGGTCTGGT GGATAAACTC
TATCATTGAT AGAGTTATGC GACCGGGGTT cgaaccccg
aataactcta-3'

VAi down (100 bases)

3'- gcttggggcc ttattgagat AGTTACTATA CGGGCCCGTA
GTTCTTCAAG CTGCAGTCTG TTCCCCCTC GCGAGGAAAA
CCGAAGGAAG GTCCGCGCCG-5'

VAi PvuII 5' (50 bases)

5'- GTGACAGCTG AGACCGTGCA AAAGGAGAGC ctgtaagcgg
gcactcttcc-3'

OVAi PvuII 5' (72 bases)

5'- GTGACAGCTG ACTCTATCAT TGATAGAGTT ATAGACCGTG
CAAAAGGAGA GCctgtaagc gggcactctt cc-3'

VAi PvuII 3' (53 bases)

3'- ccgaagggaag gtccgcgcgcg CCGACGACGC GATCGAAAAA
ACCGTCGACA GTG-5'

FIGURE 7B

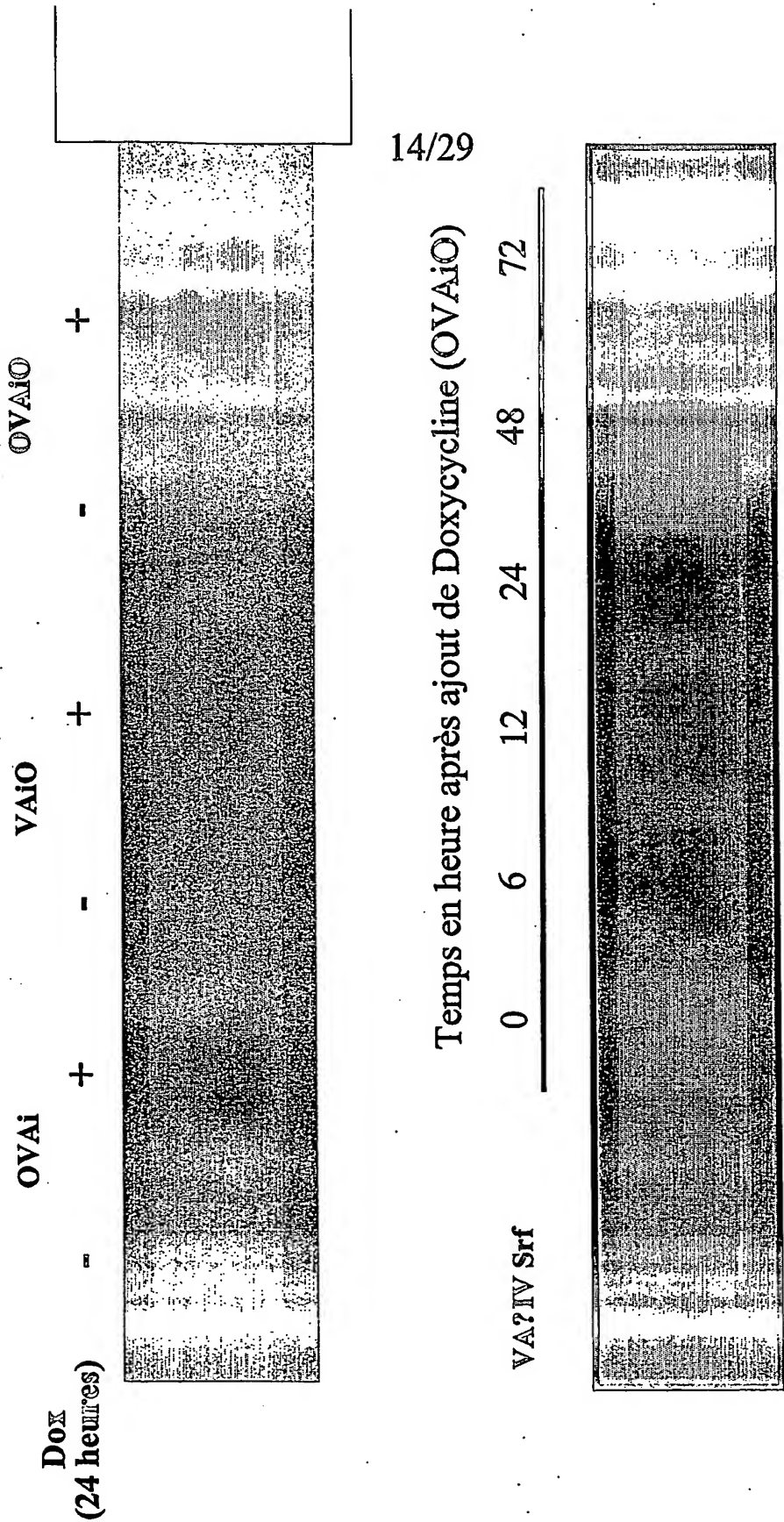


Figure 7C

15/29

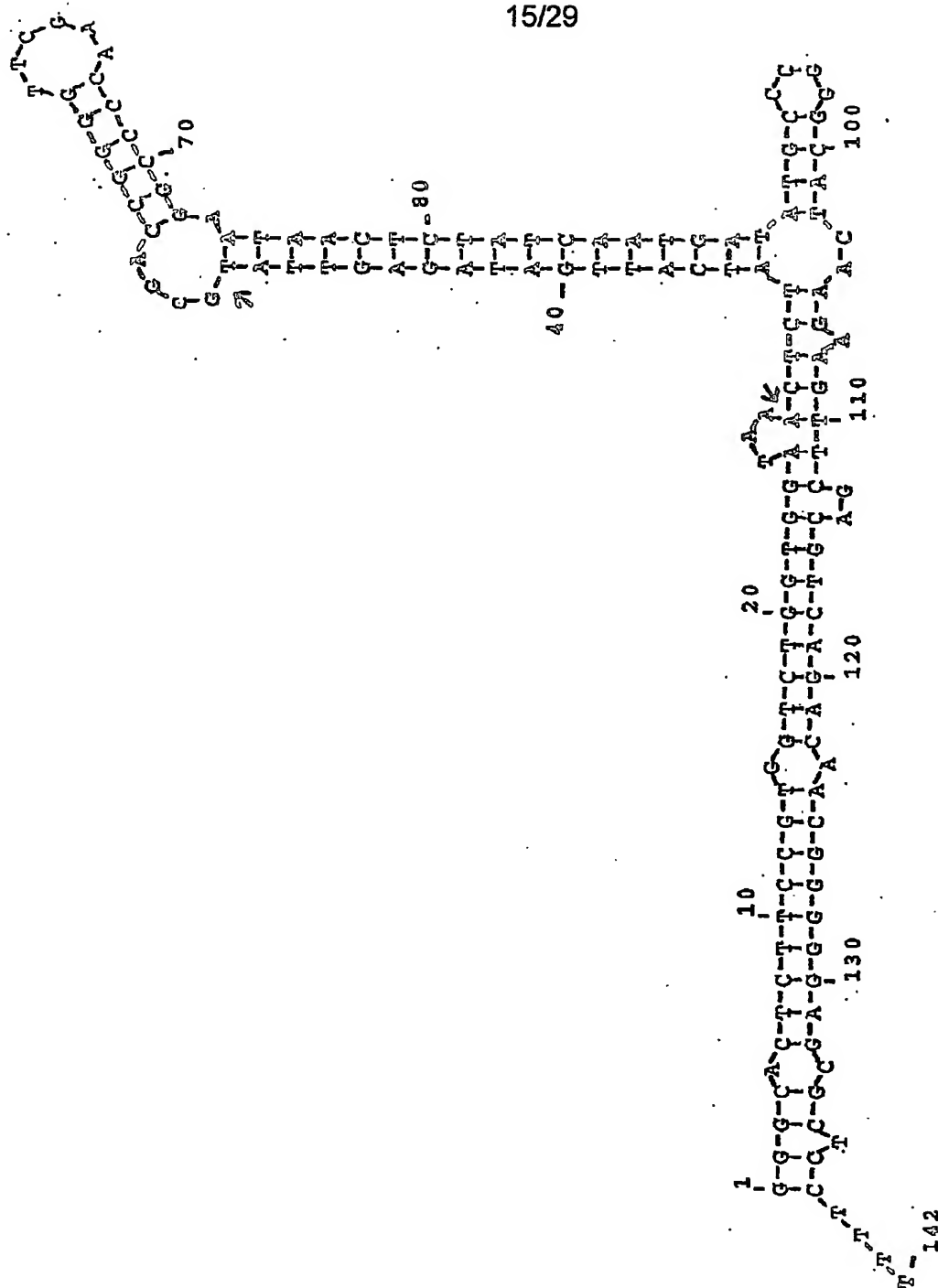


FIGURE 8 : Structure secondaire de l'ARN VA10

16/29

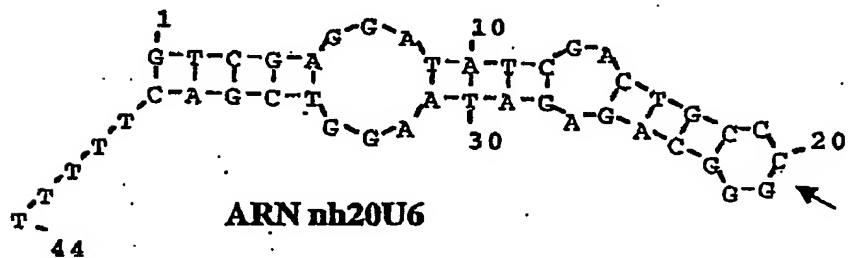
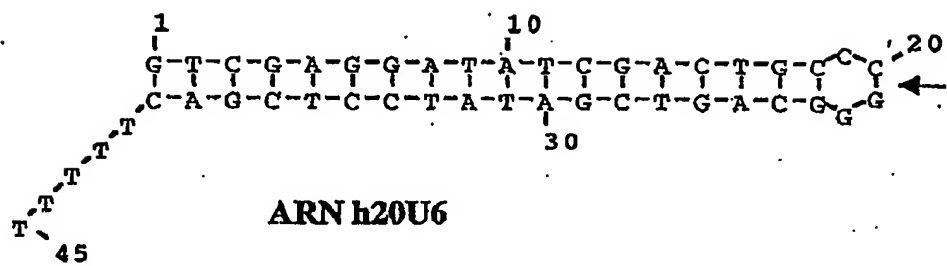
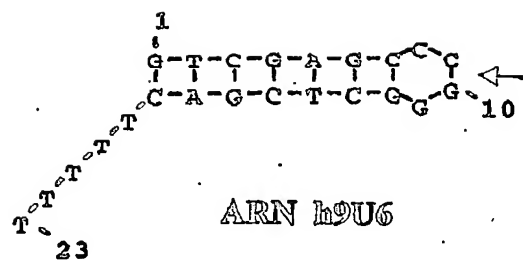


FIGURE 9 : Structure secondaire des ARN hU6

17/29

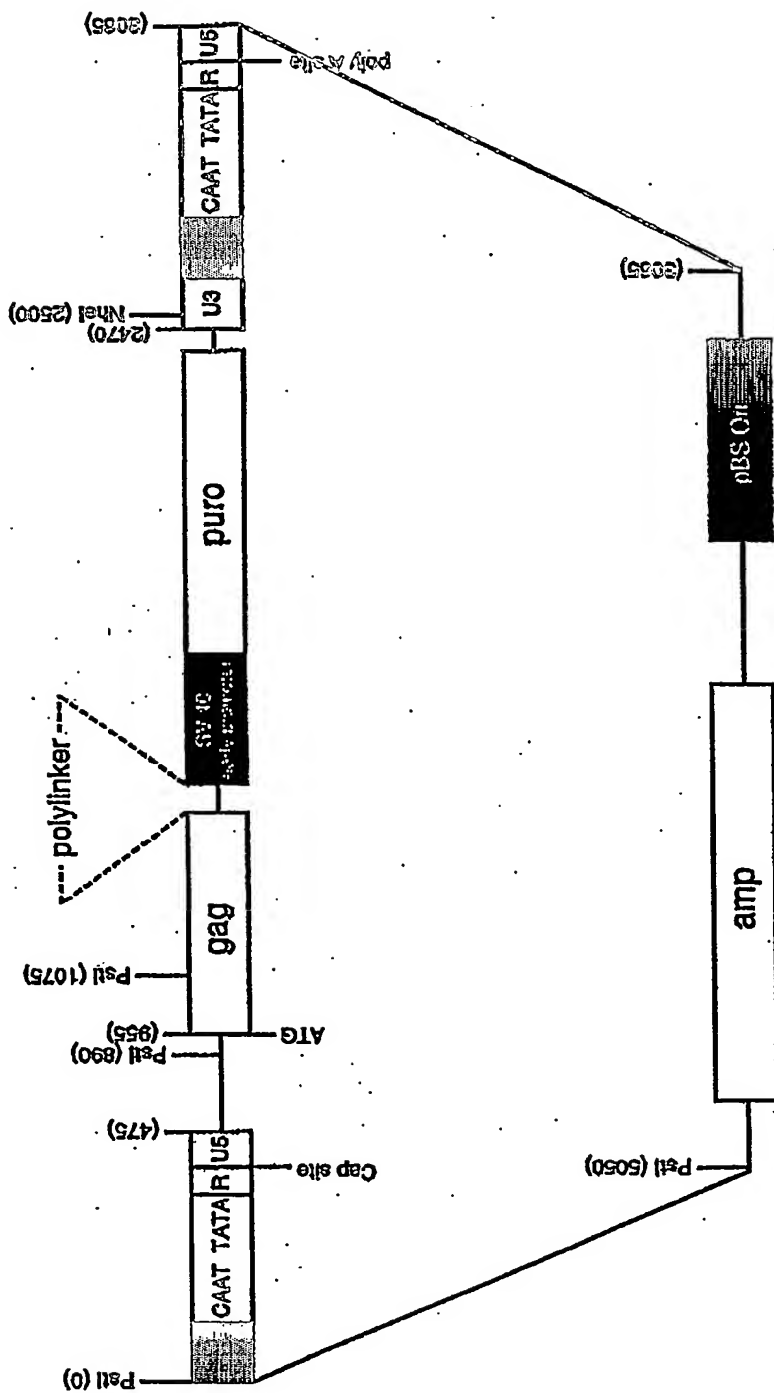


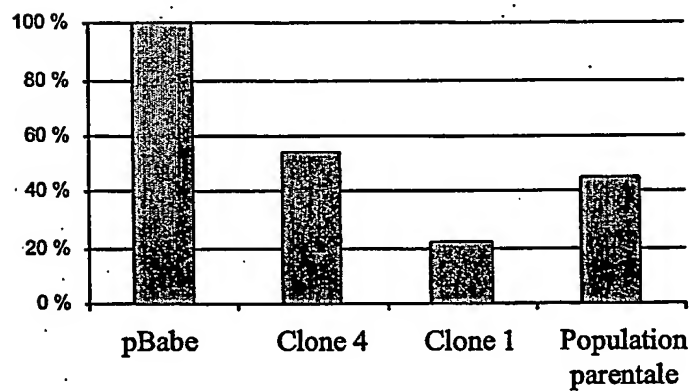
Figure 10 : plasmide pBABE puro

18/29



pBabe Clone 4 Clone 1 Population
parentale

Analyse des ARN VA TAR* par Northern-blot



Effet de l'expression des ARN VA TAR*
sur la multiplication du HIV

Figure 11

20/29

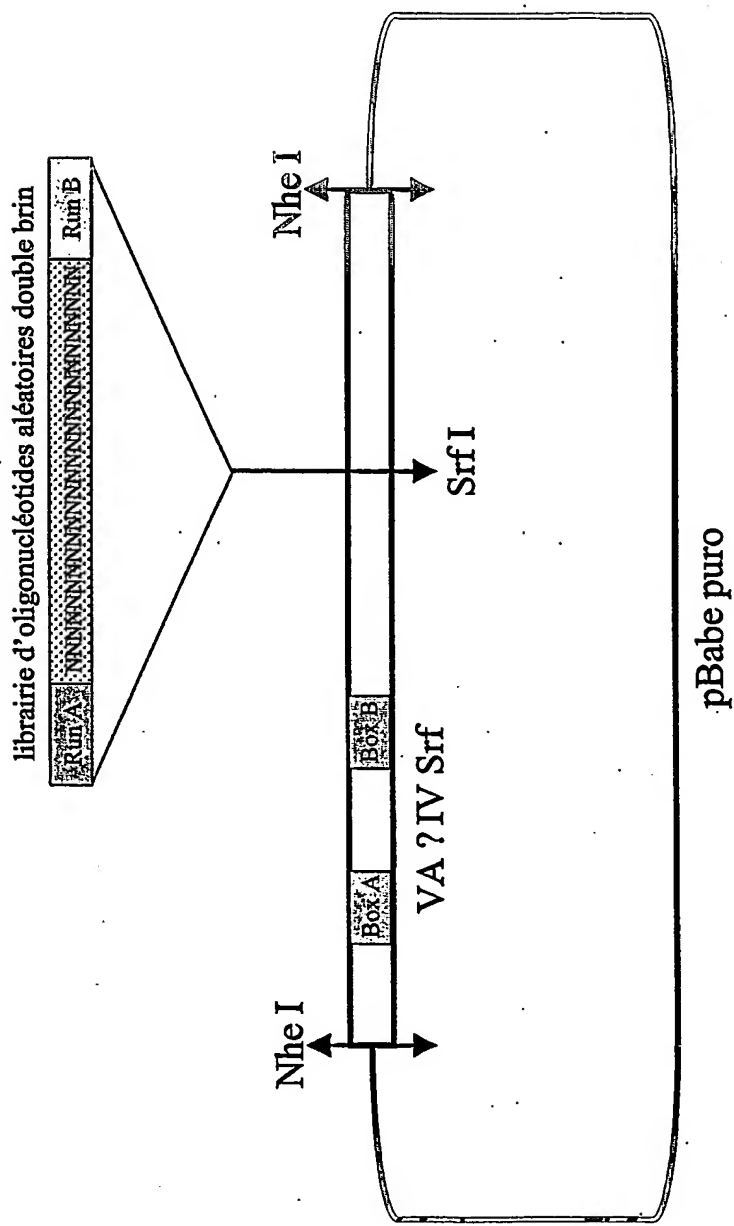


Figure 12 B

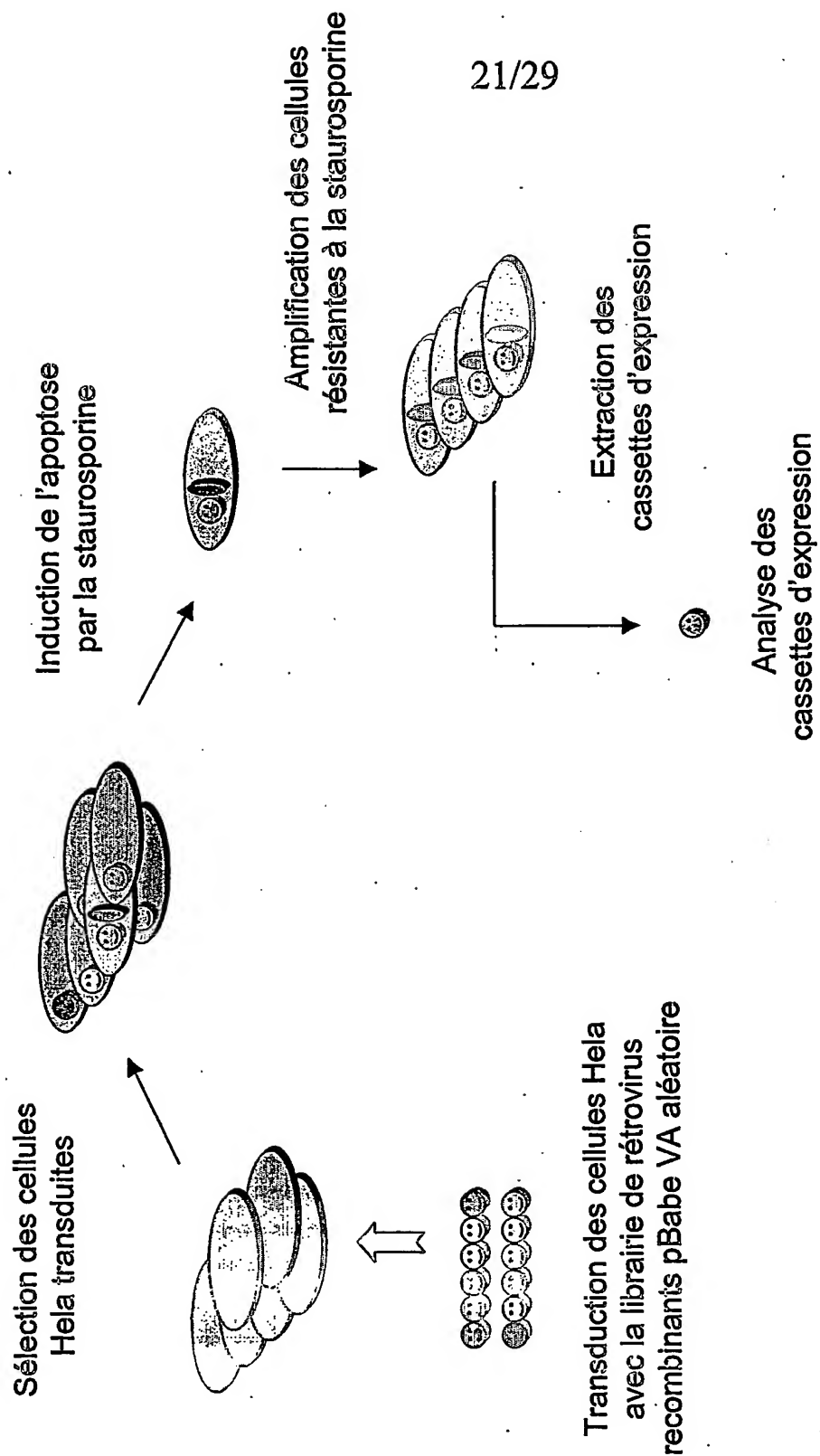
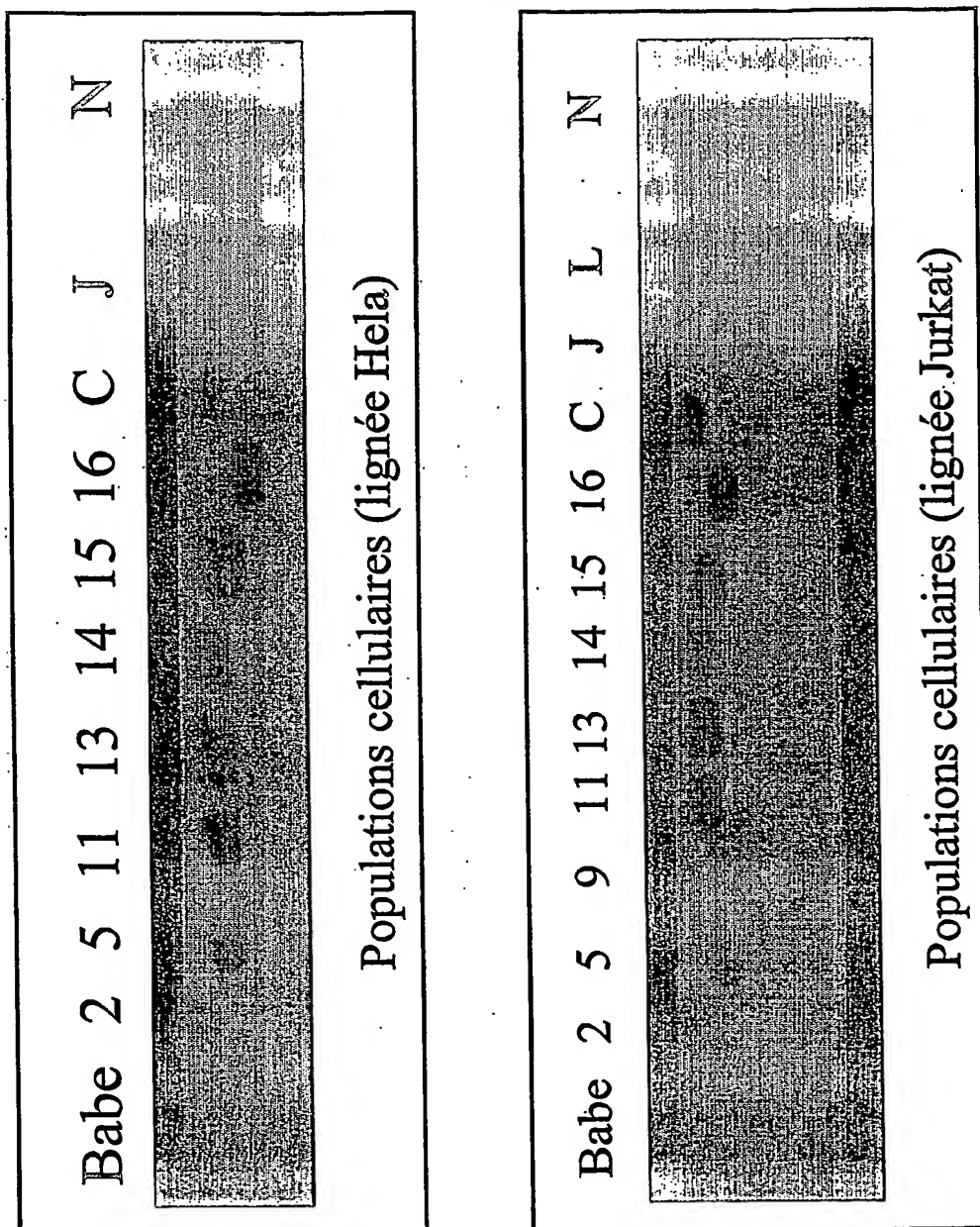


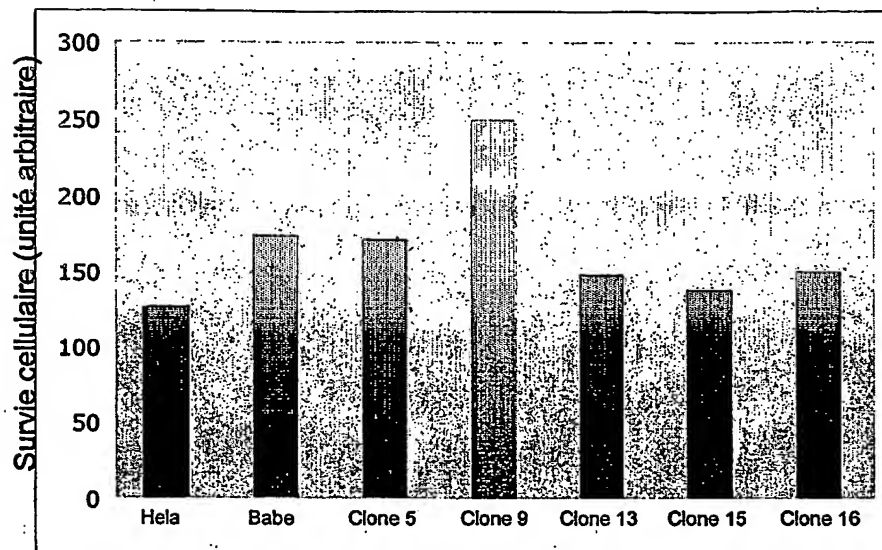
Figure 12 C

22/29

**Figure 12D**

23/29

Résistance à la Staurosporine des cellules Hela



Résistance à la Staurosporine des cellules Jurkat

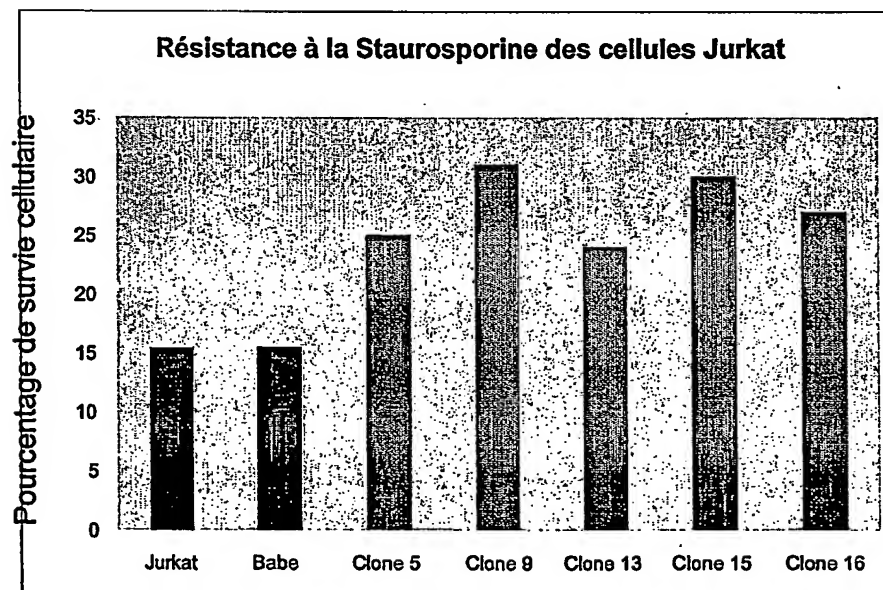
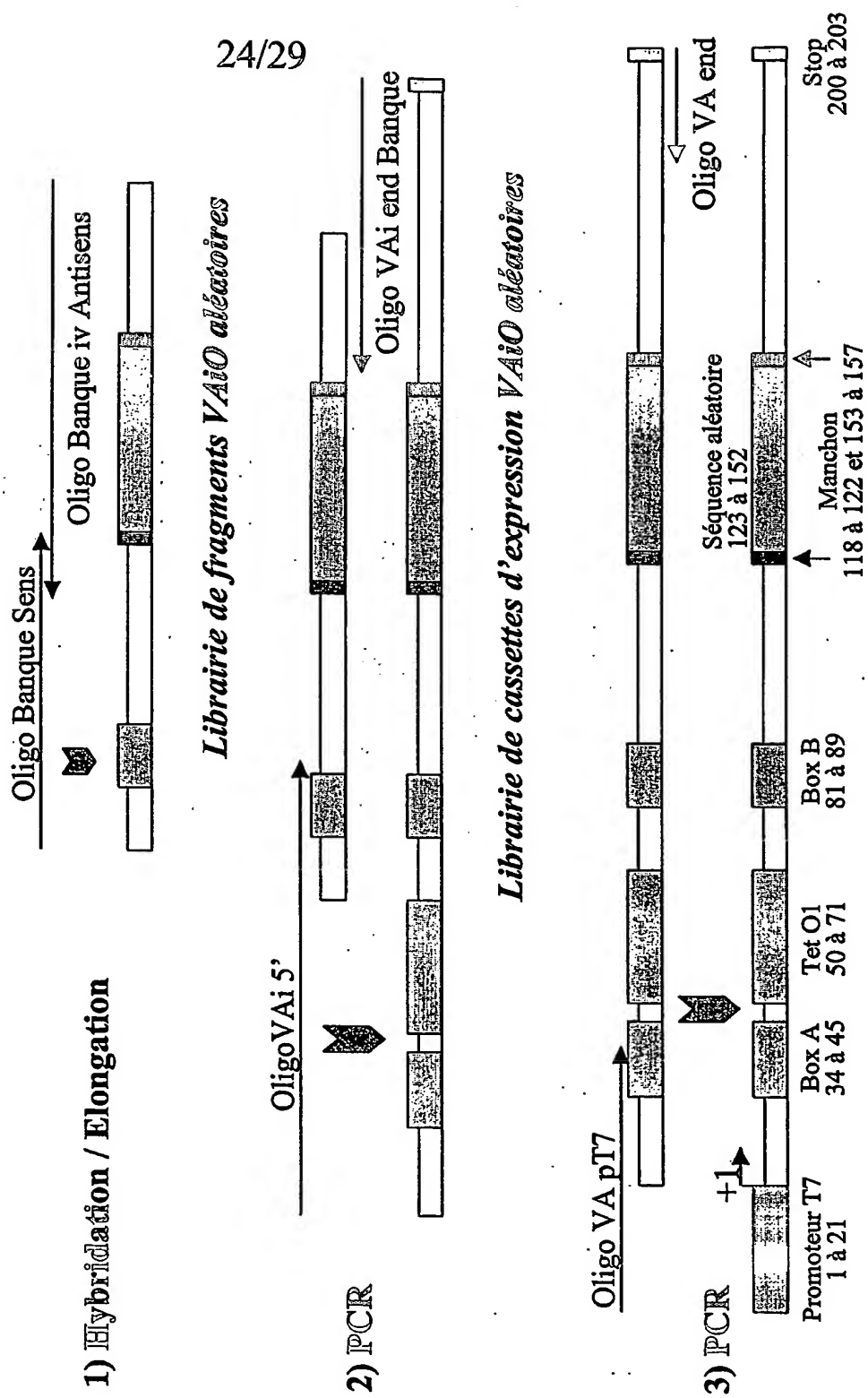


Figure 12 E



Librairie de cassettes de transcription T7 VAiO aléatoires

Figure 13 A

25/29

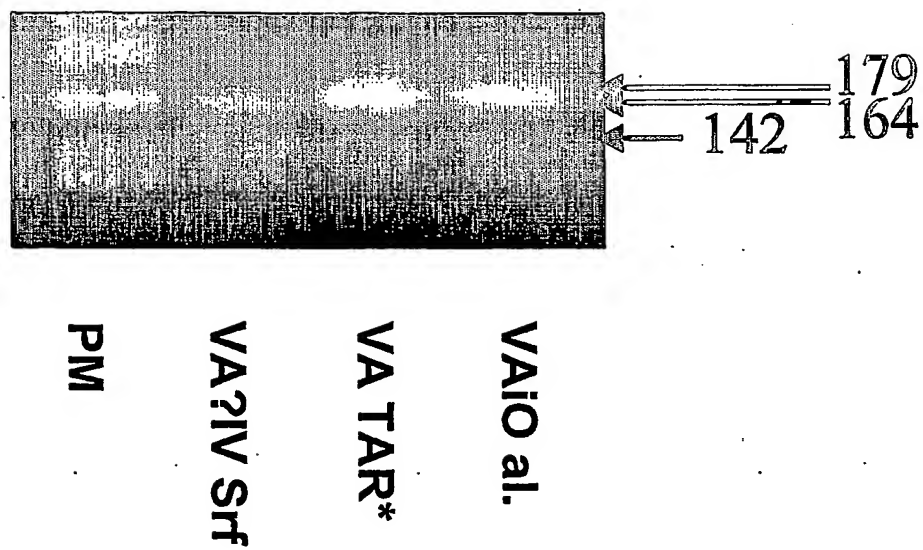


Figure 13 B

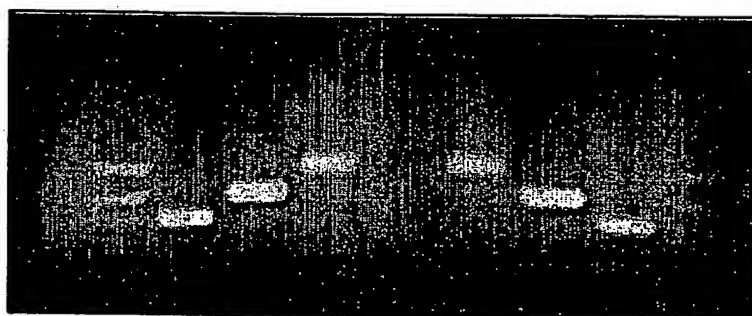
26/29

Quantité
d'ARN matrice

0

4 ng

150 ng



C-

Mélange

VA?IV Sf

VA TAR*

VAiOaléatoire

PM

VA?IV Sf

VA TAR*

VAiOaléatoire

Mélange

Figure 13 C

27/29

Etape n°1: obtention du plasmide pBabe mU6

a) Obtention du promoteur murin du gène U6

- Oligonucléotide mU6 amont

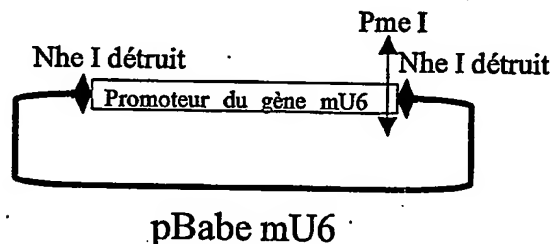
5'- CCCAAGCTTATCCGACGCCGCCATCTCTA -3'

- Oligonucléotide mU6 aval

5'- GTTTAAACAAGGCTTTTCTCCAAGGGATA - 3'

- Matrice : ADN génomique murin

b) Clonage dans le vecteur pBabe, au niveau du site NheI



Etape n°2 : obtention du plasmide pBabe mU6/VAiO

a) Obtention du gène VAiO

- Oligonucléotide VAiO Start

5'- GGGCACTCTTCCGTGGTCTGG -3'

- Oligonucléotide VAiO End NheI

5'- GCTAGCAAAAGGAGCGCTCCCCCGTTG - 3'

- Matrice : pBabe VAiO

b) Clonage dans le vecteur pBabe mU6, au niveau du site PmeI

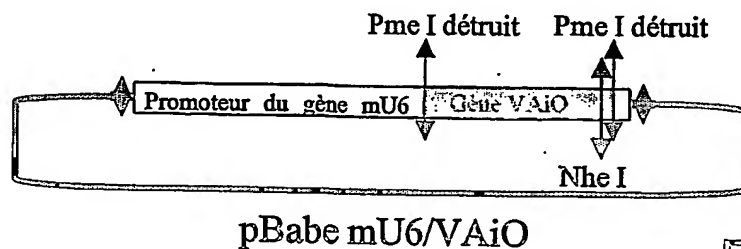


Figure 14

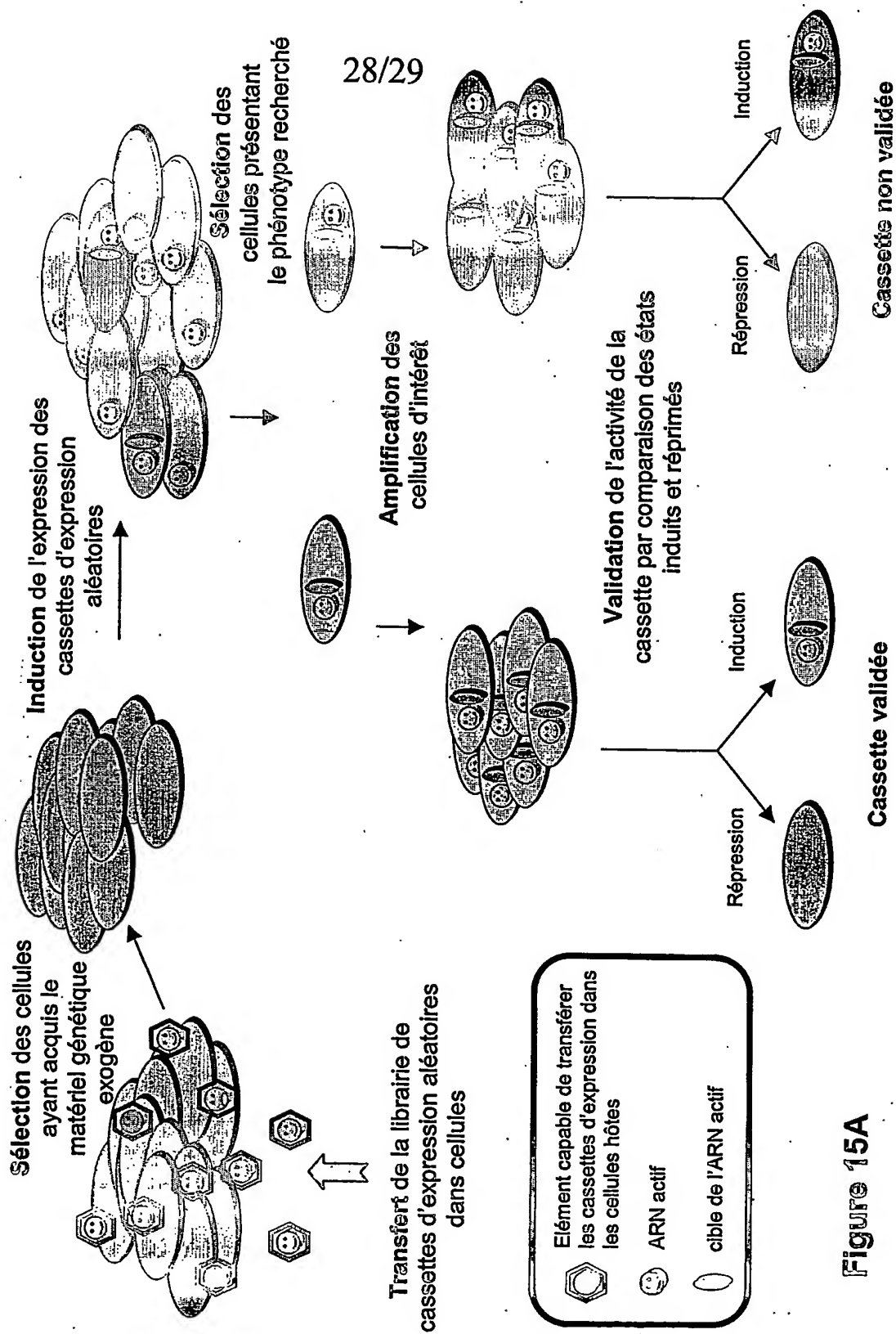


Figure 15A

29/29

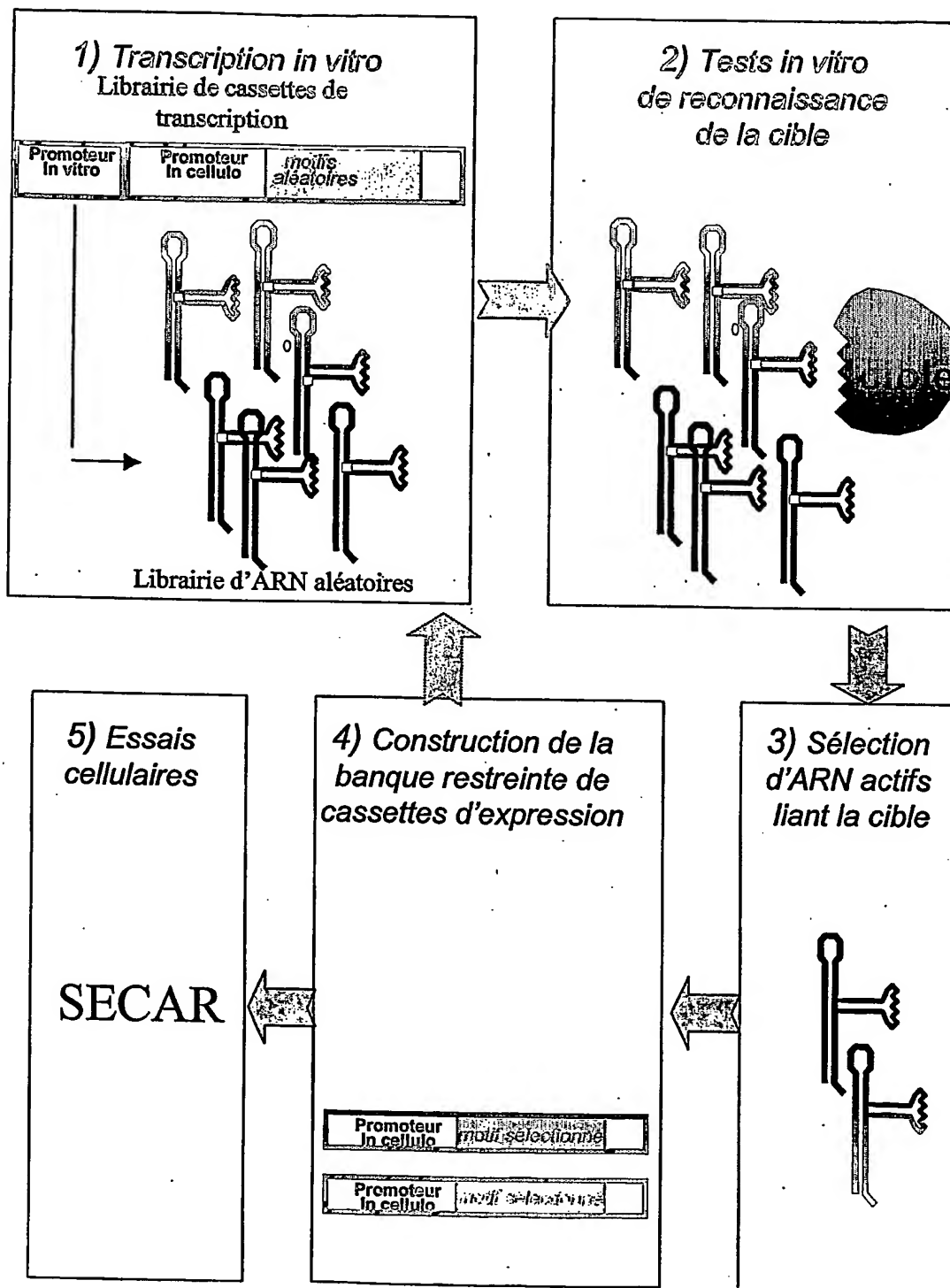


Figure 15B

B0170FR.ST25.txt
SEQUENCE LISTING

<110> université de Nice-Sophia Antipolis

<120> Méthodes et outils pour le criblage d'ARN actifs in cellulo

<130> B0170WO

<160> 23

<170> PatentIn version 3.1

<210> 1

<211> 160

<212> RNA

<213> adenovirus type 2

<400> 1
gggcacucuu ccguggucug guggauaaau ucgcaagggg aucauggcgg acgaccggg 60
uucgaacccc ggauccggcc guccgccgug auccaugcgg uuaccgccc cgugucgaac 120
ccaggugugc gacgucagac aacgggggag cgcuccuuuu 160

<210> 2

<211> 134

<212> RNA

<213> Artificial sequence

<220>

<223> modified VA1 gene

<400> 2
gggcacucuu ccguggucug guggauaaau ucgcaagggg aucauggcgg acgaccggg 60
uucgaacccc ggauccggcc guccgccgug auauccaggu gugcgacguc agacaacggg 120
ggagcgcucc uuuu 134

<210> 3

<211> 142

B0170FR.ST25.txt

<212> RNA

<213> Artificial Sequence

<220>

<223> modified VA1 gene

<400> 3

gggcacucuu ccguggucug guggauaaa ucgcaagggg aucauggcgg acgaccgggg 60

uucgaacccc ggauccggcc guccgccgug augcccgggc auccaggugu gcgacgucag 120

acaacggggg agcguccuu uu 142

<210> 4

<211> 141

<212> RNA

<213> Artificial Sequence

<220>

<223> modified VA1 gene

<400> 4

gggcacucuu ccguggucug guggauaaa ucgcaagggg aucauggcgg acgaccgggg 60

uucgaacccc ggauccggcc guccgccgug augcccgggc auccaggugu gcgacgucaa 120

uaaacggggg agcgccuuu u 141

<210> 5

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> modification of VA1 gene

<400> 5

gcccgggcat ccaggtgtgc gacgtcaata aacgggggag cgccctttt 49

<210> 6

<211> 23

<212> DNA

<213> Artificial Sequence

B0170FR.ST25.txt

<220>

<223> h9U6

<400> 6

tcgagcccgg gctcgacttt ttc

23

<210> 7

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> h9U6

<400> 7

gtcgagcccgg ggctcgactt ttt

23

<210> 8

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> h20U6

<400> 8

tcgaggatat cgactgcccgg ggcagtcgat atcctcgact ttttc

45

<210> 9

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> h20U6

<400> 9

gtcgaggata tcgactgccc ggcagtcga ttcctcgac ttttt

45

<210> 10

<211> 44

<212> DNA

B0170FR.ST25.txt

<213> Artificial Sequence

<220>

<223> nh20U6

<400> 10

tcgaggatat cgactgcccg ggcagagata aggtcgactt tttc

44

<210> 11

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> nh20U6

<400> 11

gtcgaggata tcgactgccc ggcagagat aaggtcgact tttt

44

<210> 12

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> U6 sfr

<400> 12

gtgggcatg ggtgcccggg ctttcgtcct ttccacaag

39

<210> 13

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> U6Tt

<400> 13

gcggacttcg gtccgctttt t

21

<210> 14

B0170FR.ST25.txt.

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> U6Tt

<400> 14

gtcgacccat gctagagcgg acttcggtcc gcttttt

37

<210> 15

<211> 142

<212> DNA

<213> Artificial Sequence

<220>

<223> VAiO gene

<400> 15

gggcactctt ccgtggtctg gtggataaac tctatcattg atagagttat gcgaccggg 60

ttcgaacccc ggaataactc tatcaatgat atgccccggc atcaagaagt tcgacgtcag 120

acaacggggg agcgtcctt tt 142

<210> 16

<211> 191

<212> DNA

<213> Artificial Sequence

<220>

<223> OVAi gene

<400> 16

actctatcat tgatagagtt atagaccgtg caaaaggaga gcctgtaagc gggcactctt 60

ccgtggtctg gtggataaac tctatcattg tagagttatg cgaccggggt tcgaacccc 120

gaataactct atcaatgata tgccccggca tcaagaagtt cgacgtcaga caacggggga 180

gcgctcctt t 191

<210> 17

<211> 192

<212> DNA

B0170FR.ST25.txt

<213> Artificial Sequence

<220>

<223> OVAiO gene

<400> 17

actctatcat tgatagagtt atagaccgtg caaaaggaga gcctgtaagc gggcactctt 60

ccgtgggtctg gtggataaac tctatcattg atagagttat gcgaccgggg ttcgaacccc 120

ggaataactc tatcaatgat atgcccgggc atcaagaagt tcgacgttag acaacggggg 180

agcgctcctt tt 192

<210> 18

<211> 90

<212> DNA

<213> Artificial Sequence

<220>

<223> primer VAI up

<400> 18

ctgtaagcgg gcactcttcc gtggtctggt ggataaactc tatcattgat agagttatgc 60

gaccgggggtt cgaaccccg aataactcta 90

<210> 19

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> primer VAI down

<400> 19

gcttggggcc ttattgagat agttactata cgggcccgta gttcttcaag ctgcagtctg 60

ttgccccctc gcgaggaaaa ccgaagggaag gtccgcgccg 100

<210> 20

<211> 50

<212> DNA

<213> Artificial Sequence

B0170FR.ST25.txt

<220>

<223> primer VAI PvuII 5'

<400> 20

gtgacagctg agaccgtgca aaaggagagc ctgtaagcgg gcactcttcc 50

<210> 21

<211> 72

<212> DNA

<213> Artificial Sequence

<220>

<223> primer OVAi PvuII 5'

<400> 21

gtgacagctg actctatcat tgatagagtt atagaccgtg caaaaggaga gcctgtaagc 60

gggcactctt cc 72

<210> 22

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> primer VAI PvuII 3'

<400> 22

ccgaagggaag gtccgcgccg ccgacgacgc gatcgaaaaa accgtcgaca gtg 53

<210> 23

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> sequence TAR

<400> 23

acgaagcttg atcccgtatg ccggtcgatc gcttcga 37

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.